

TESIS DOCTORAL

EVALUACIÓN DE RIESGOS PSICOSOCIALES CON MEDIDAS DE AUTOINFORME. DIFICULTADES Y AMENAZAS A LA VALIDEZ

M^a Rita Louzán Mariño

PSICOLOXÍA SOCIAL, BÁSICA E METODOLOXÍA - PROGRAMA
DE DOUTORAMENTO EN PROCESOS PSICOLÓXICOS E
COMPORTAMENTO SOCIAL-FACULTADE DE PSICOLOXÍA

SANTIAGO DE COMPOSTELA

2017



TESIS DOCTORAL

EVALUACIÓN DE RIESGOS PSICOSOCIALES CON MEDIDAS DE AUTOINFORME. DIFICULTADES Y AMENAZAS A LA VALIDEZ

Fdo:

M^a Rita Louzán Mariño

PSICOLOXÍA SOCIAL, BÁSICA E METODOLOXÍA - PROGRAMA
DE DOUTORAMENTO EN PROCESOS PSICOLÓXICOS E
COMPORTAMENTO SOCIAL-FACULTADE DE PSICOLOXÍA

SANTIAGO DE COMPOSTELA

2017



INFORME DOS DIRECTORES DA TESE

D. DÁMASO RODRÍGUEZ FERNÁNDEZ, Profesor do Departamento de Psicoloxía Social, Básica e Metodoloxía da Universidade de Santiago de Compostela e D. CARLOS MONTES PIÑEIRO, Profesor da Escola Universitaria de Relacións Laborais adscrita á Universidade de Santiago de Compostela.

Como Directores da Tese de Doutoramento titulada "AVALIACIÓN DE RISCOS PSICOSOCIAIS CON MEDIDAS DE AUTOINFORME: DIFICULTADES E AMEAZAS Á VALIDEZ"

Presentada por DONA MARÍA RITA LOUZÁN MARÍÑO, alumna do Programa de Doutoramento en Procesos Psicolóxicos e Comportamento Social

Autorizan a presentación da tese indicada, considerando que reúne os requisitos esixidos no artigo 34 do regulamento de Estudos de Doutoramento, e que como Directores da mesma non incorren nas causas de abstención establecidas na lei 40/2015.

En Santiago de Compostela, a 15 de setembro de 2017.

Asdo.:

D. Dámaso Rodríguez Fernández

Asdo:

D. Carlos Montes Piñeiro



*Dedicado a mi familia, especialmente a la memoria de mi padre.
Su ausencia ha convertido este proyecto de investigación en mi refugio.*





*Considero más valiente al que conquista sus deseos
que al que conquista a sus enemigos, ya que la
victoria más dura es la victoria sobre uno mismo.*

Aristóteles (384 a.C.-322 a.C.) Filósofo griego.





AGRADECIMIENTOS

Quiero expresar un especial agradecimiento a todas aquellas personas, que directa o indirectamente, de modo personal o académicamente, han contribuido a que este trabajo llegue a su fin, entre ellos:

Los 16 expertos encuestados por su desinteresada y valiosa participación en el estudio Delphi. Sin su tiempo y compromiso este proyecto no hubiese podido realizarse: Josep M. Blanch Ribas, Eduardo García Cueto, Pedro R. Gil Monte, Joan Guàrdia Olmos, Daniel Iglesias Pastrana, María José Lasa Gorraiz, Francisco López Barón, Javier Llana Álvarez, Jenny M^a Manteiga Rilo, Albert Mariné, Margarita Oncins de Frutos, Gabriel Pérez Zambrana, Mari Carmen Ramos Lema, Marisa Salanova Soria, Guillermo Soriano Tarín y Placer Vieco Marcos.

Los centenares de trabajadores anónimos que han querido contribuir a esta investigación y aportar su granito de arena rellenando los cuestionarios, también de manera desinteresada.

Los directores de esta tesis: Dámaso Rodríguez y Carlos Montes por sus consejos, tiempo y dedicación.

Alberto Villarroya por todo el apoyo moral y paciencia en los momentos difíciles, que los hubo.

Todas las personas que con sus palabras o silencios me han infundido ánimo.



Evaluación de riesgos psicosociales con medidas de autoinforme. Dificultades y amenazas a la validez.

RESUMO

A pesar das repercusións dos riscos psicosociais na saúde dos traballadores, a miúdo son subestimados polas organizacións, alegando que son factores de difícil xestión. Os motivos de que isto ocorra son diversos, entre eles: unha pobre cultura preventiva, a falta de lexislación específica ou o mal uso dos instrumentos de medida. Concretamente, unha das principais críticas á Ergonomía e á Psicosocioloxía aplicada céntrase no uso de medidas de autoinforme como medio de obtención de información. Cuestionase a súa validez argumentando que son susceptibles a diversas fontes de distorsión como os sesgos de resposta ou á afectividade negativa. No intento de mellorar a calidade das avaliacións de riscos psicosociais, realizáronse dous estudos. En primeiro lugar, un cualitativo utilizando o método Delphi, onde se convidou a participar a un elenco de reputados Psicólogos do traballo e Técnicos de prevención de riscos laborais, que de modo consensuado, sacaron á luz as principais dificultades e ameazas á validez que entorpecen as avaliacións de riscos psicosociais. Á súa vez recomendaron unha serie de procedementos e actuacións para a mellora das mesmas. O segundo e último dos estudos é de corte cuantitativo e achega evidencias de distorsión do sesgo de deseabilidade social e da variable confusora de afecto negativo na escala CoPsoQ-ISTAS 21 v.2., ofrecendo técnicas estatísticas específicas e detalladas para o seu control. As conclusións de ambos os estudos contribúen a mellorar a calidade e validez das avaliacións de risco psicosocial de orixe laboral, sendo de gran utilidade para a práctica profesional.

PALABRAS CHAVE

Avaliación de riscos psicosociais, ameazas á validez, deseabilidade social, afectividade negativa, Delphi.

RESUMEN

A pesar de las repercusiones de los riesgos psicosociales en la salud de los trabajadores, a menudo son subestimados por las organizaciones, alegando que son factores de difícil gestión. Los motivos de que esto ocurra son diversos, entre ellos: una pobre cultura preventiva, la falta de legislación específica o el mal uso de los instrumentos de medida. Concretamente, una de las principales críticas a la Ergonomía y a la Psicosociología Aplicada se centra en el uso de medidas de autoinforme como medio de obtención de información. Se cuestiona su validez argumentando que son susceptibles a diversas fuentes de distorsión como los sesgos de respuesta o la afectividad negativa. En el intento de mejorar la calidad de las evaluaciones de riesgos psicosociales, se han realizado dos estudios. En primer lugar, uno cualitativo utilizando el método Delphi, donde se ha invitado a participar a un elenco de reputados Psicólogos del trabajo y Técnicos de prevención de riesgos laborales, que de modo consensuado, han sacado a la luz las principales dificultades y amenazas a la validez que entorpecen las evaluaciones de riesgos psicosociales. A su vez han recomendado una serie de procedimientos y actuaciones para la

mejora de las mismas. El segundo y último de los estudios es de corte cuantitativo y aporta evidencias de distorsión del sesgo de deseabilidad social y de la variable confusora de afecto negativo en la escala CoPsoQ-ISTAS 21 v.2., ofreciendo técnicas estadísticas específicas y detalladas para su control. Las conclusiones de ambos estudios contribuyen a mejorar la calidad y validez de las evaluaciones de riesgo psicosocial de origen laboral siendo de gran utilidad para la práctica profesional.

PALABRAS CLAVE

Evaluación de riesgos psicosociales, amenazas a la validez, deseabilidad social, afectividad negativa, Delphi.

SUMMARY

Despite the repercussion of psychosocial risks on workers' health, they are often underestimated by organizations, arguing that they are difficult to manage. The reasons for this are diverse, including poor preventive culture, lack of specific legislation or misuse of measuring instruments. Specifically, one of the main criticisms of Ergonomics and Applied Psychosociology focuses on the use of self-report measures as a mean of obtaining information. Their validity is questioned, arguing that they are susceptible to different sources of distortion such as response bias or negative affectivity. In order to improve the quality of psychosocial risk assessments, two studies have been carried out. First, a qualitative one using the Delphi method, where a group of well-known Work Psychologists and Occupational Risk Prevention Technicians have been invited to participate, which have brought to light in a consensual way the main difficulties and threats to the validity that hamper psychosocial risk assessments. At the same time, they have recommended a number of procedures and actions to improve them. The second and last of the studies is quantitative and provides evidence of distortion of the social desirability bias and the confounding variable of negative affectivity in the CoPsoQ-ISTAS 21 v.2 scale, offering specific and detailed statistical techniques for its control. The conclusions of both studies contribute to improve the quality and validity of work-related psychosocial risk assessments, being very useful for professional practice.

KEYWORDS

Psychosocial risk assessments, threats to validity, social desirability, negative affectivity, Delphi.

INDICE

INTRODUCCION	1
1. FACTORES DE RIESGO PSICOSOCIAL VS RIESGOS PSICOSOCIALES EN EL TRABAJO	11
1.1. FACTORES PSICOSOCIALES DE RIESGO.	11
1.2. RIESGO PSICOSOCIAL	13
1.3. LOS FACTORES PSICOSOCIALES Y SU RELACIÓN CON LA SALUD	15
1.4. PRINCIPALES RIESGOS PSICOSOCIALES EN EL TRABAJO Y SU RELACIÓN CON LA SALUD	17
1.4.1. Estrés laboral.	17
1.4.1.1. Delimitación conceptual	17
1.4.1.2. El estrés laboral como precursor de enfermedad	19
1.4.2. Síndrome de <i>Burnout</i>	23
1.4.3. Mobbing	26
1.5. PANORAMA ACTUAL Y PERSPECTIVAS DE FUTURO DE LOS RIESGOS PSICOSOCIALES	29
2. EVALUACION DE RIESGOS PSICOSOCIALES: INSTRUMENTOS DE MEDIDA	35
2.1. EVALUACIÓN DE RIESGOS PSICOSOCIALES. MARCO LEGAL	35
2.2. METODOLOGÍAS Y TÉCNICAS DE EVALUACIÓN DE FACTORES DE ORIGEN PSICOSOCIAL EN EL TRABAJO	36
2.3. CRITERIOS PARA SELECCIONAR LA HERRAMIENTA DE EVALUACIÓN IDÓNEA ATENDIENDO A SU DISEÑO	39
2.3.1. Modelos Teóricos explicativos del estrés.	40
2.3.1.1. Modelo Demanda-Control	41
2.3.1.2. Modelo Esfuerzo-Recompensa (ER)	42
2.3.2. Medición de los factores psicosociales de riesgo	43
2.3.3. Test Psicométricos: Conceptualización y clasificación	44
2.3.4. Rigor científico en los Test	45
2.3.4.1. Fiabilidad	46
2.3.4.2. Validez	47
2.3.4.2.1. Validez de contenido	48
2.3.4.2.2. Validez de constructo	48
2.3.4.2.3. Validez de criterio o predictiva	48
2.3.5. Cuestionario de autoinforme aplicados a la evaluación de riesgos psicosociales.	48
2.4. REVISIÓN DE LAS PRINCIPALES TÉCNICAS BASADAS EN ESCALAS AUTOINFORME UTILIZADAS EN LA EVALUACIÓN DE RIESGOS PSICOSOCIALES	50

3. APLICACIÓN, CORRECCION E INTERPRETACION DE RESULTADOS DE LAS MEDIDAS DE AUTOINFORME	65
3.1. AMENAZAS A LA VALIDEZ EN LAS MEDIDAS DE AUTOINFORME	65
3.2. SESGOS DE RESPUESTA	72
3.2.1. Tendencia de respuesta: Deseabilidad Social	74
3.2.2. Estilos de respuesta: Respuesta Aquiescente	77
3.3. CONTROL DE SESGOS	79
3.3.1. Formas de control de la deseabilidad social en medidas de autoinforme	80
3.3.1.1. Escalas específicas	80
3.3.1.2. Otros métodos de control	81
3.3.2. Formas de control de aquiescencia en medidas de autoinforme	82
3.3.2.1. Escalas específicas	82
3.3.2.2. Otros métodos de control	83
3.3.2.2.1. <i>Escalas balanceadas</i>	83
3.3.2.2.2. <i>Escalas con pares de ítems de significado opuesto</i>	83
3.3.2.2.3. <i>Escalas con test ipsativos</i>	84
3.3.2.2.4. <i>Otras recomendaciones para evitar la aquiescencia en los autoinformes</i>	84
3.4. SESGO DE VARIANZA COMÚN DEL METODO	85
4. INVESTIGACION CUALITATIVA: METODO DELPHI	91
4.1. INTRODUCCIÓN	91
4.1.1. Antecedentes del método Delphi	91
4.1.2. Características principales	93
4.2. JUSTIFICACIÓN DEL PROBLEMA Y OBJETIVOS DE LA INVESTIGACIÓN	94
4.2.1. Objetivos	96
4.3. METODOLOGÍA DE ESTUDIO	96
4.3.1. Formulación del problema	97
4.3.2. Selección de expertos	99
4.3.2.1. Enfoque cuantitativo: Número de rondas y tamaño de la muestra.	100
4.3.2.2. Enfoque cualitativo: Criterio de Selección de expertos	101
4.3.1. Control de errores y sesgos	103
4.3.2. Gestión 1ª ola	103
4.3.2.1. Confección y envío de cuestionario	103
4.3.2.2. Valoraciones de expertos y tratamiento estadístico	105
4.3.2.3. Análisis estadístico 1ª ronda: Posición, tendencia central, consenso y dispersión	105
4.3.2.4. Criterios de decisión de consenso	107
4.3.3. Gestión 2ª ola	108
4.3.3.1. Confección y envío del cuestionario	108
4.3.3.2. Análisis estadístico y criterios de consenso 2ª ronda.	109

4.3.4. Estabilidad entre rondas	109
4.3.4.1. Variación de las medidas de dispersión	111
4.3.4.2. Prueba de Wilcoxon.	112
4.3.4.3. Coeficiente Kappa de Cohen	112
4.4. RESULTADOS	112
4.4.1. Resultados 1ª ronda ítems cuantitativos	112
4.4.2. Resultados 2ª ronda ítems cuantitativos	114
4.4.3. Resultados 2ª ronda ítems cualitativos	118
4.4.3.1. Bloque 1: Respuestas a las preguntas abiertas	118
4.4.3.2. Bloque 2: Listado de comentarios y sugerencias	120
4.4.4. Resultados estabilidad individual	123
4.4.4.1. La proporción de expertos que modifican su valoración	124
4.4.5. Resultados de estabilidad entre rondas	124
4.5. DISCUSIÓN DE LOS RESULTADOS Y CONCLUSIONES	126
4.5.1. Ítems consensuados	127
4.5.1.1. Apartado de afirmaciones generales (Ítems A1-A4)	127
4.5.1.2. Apartado de diseño de instrumentos de medida (Ítems A5-A11)	128
4.5.1.3. Apartado de aplicación (ítems A12-A18)	131
4.5.1.4. Apartado de corrección e interpretación de los resultados (Ítems A19-A25)	132
4.5.1.5. Apartado de sugerencias y observaciones	133
4.5.2. Ítems sin consenso	133
4.5.3. Implicaciones teóricas y prácticas	135
4.5.4. Limitaciones y líneas de investigación futuras	137
5. INVESTIGACIÓN CUANTITATIVA	141
5.1. INTRODUCCIÓN	141
5.1.1. Deseabilidad social como distorsión de las respuestas	141
5.1.1. Afecto negativo como variable de confusión	142
5.2. JUSTIFICACIÓN DEL PROBLEMA, OBJETIVOS E HIPÓTESIS DE LA INVESTIGACIÓN	144
5.2.1. Objetivo general	144
5.2.2. Objetivos específicos	144
5.2.3. Hipótesis	144
5.3. METODOLOGÍA DE ESTUDIO	145
5.3.1. Selección de herramienta de evaluación de riesgos psicosociales	145
5.3.2. Diseño y Procedimiento de recogida de datos	146
5.3.3. Depuración de datos	148
5.3.4. Participantes	148
5.3.4.1. Distribución por sexo y edad	149

5.3.4.2. Distribución en función de la nacionalidad	150
5.3.4.3. Distribución según la situación laboral	150
5.3.4.4. Distribución según el puesto de trabajo que ocupan	151
5.3.4.5. Distribución según antigüedad en el puesto de trabajo y en la empresa	151
5.3.4.6. Distribución según tiempo de baja médica en los últimos 12 meses	152
5.3.5. Instrumentos	152
5.3.5.1. Cuestionario de Evaluación de Riesgos Psicosociales en el Trabajo del método ISTAS21 (CoPsoQ) (versión 2)	153
5.3.5.2. Escala de afecto positivo y negativo (PANAS)	156
5.3.5.3. Ítems de control de la deseabilidad social	157
5.3.5.4. Fiabilidad de las escalas de medida empleadas	157
5.3.6. Análisis de datos	158
5.3.6.1. Sesgo de varianza común del método	159
5.3.6.2. Efecto de la Deseabilidad Social en la escala CoPsoQ-ISTAS 21 v.2	159
5.3.6.3. Funcionamiento Diferencial de Ítem.	160
5.4. RESULTADOS	163
5.4.1. Exploración de la varianza común de método	163
5.4.2. Análisis descriptivos	163
5.4.3. Depuración del sesgo de Deseabilidad social	165
5.4.4. Funcionamiento Diferencial del ítem según el AN:	171
5.5. DISCUSIÓN DE LOS RESULTADOS Y CONCLUSIONES	174
5.5.1. Discusión de resultados y conclusiones sobre el control de la DS	174
5.5.2. Discusión resultados y conclusiones sobre la detección de DIF	176
5.5.3. Implicaciones teóricas y prácticas	178
5.5.4. Limitaciones y líneas de investigación futuras	179
6. CONCLUSIONES	183
6.1. CONCLUSIONES FINALES	183
REFERENCIAS	191
INDICE TABLAS	303
INDICE FIGURAS	305
ANEXO I	307
ANEXO II	313

INTRODUCCION

Los riesgos psicosociales constituyen una seria amenaza para la salud de los trabajadores, y tanto es así que se ha catalogado el estrés como la *epidemia laboral del siglo XXI* (Baratech, 2006). Tal enunciado no debe de considerarse baladí, pues a comienzos de este siglo uno de cada 3 trabajadores europeos, aproximadamente 40 millones de personas, experimentaba los efectos nocivos del estrés derivado de las condiciones de trabajo (Eurofound, 2007). Lejos de mejorar con el paso de los años, las cifras se han agravado a consecuencia del azote de la crisis económica mundial (Regus, 2012), que ha lesionado profundamente el panorama laboral debido a las políticas de ajuste de las empresas. Éstas han actuado en menoscabo de las condiciones laborales de los trabajadores, dando lugar a despidos masivos en grandes empresas, principalmente en la construcción y en el sector industrial (Luceño, Martín, Rubio y Jaén, 2008), incrementándose a su vez el empleo informal y la inmigración irregular (Porthé et al 2009). Esta realidad puede verse reflejada en el Sondeo de Opinión Paneuropeo Sobre Seguridad y Salud en el Trabajo elaborado por la Agencia Europea para la Seguridad y Salud Laboral (EU-OSHA, 2013), donde un 72% de trabajadores considera la inseguridad laboral y la reorganización del trabajo como las principales causas del estrés.

A pesar de haber transcurrido casi una década desde el comienzo de la crisis, no todos los países europeos han seguido el mismo ritmo de recuperación. Así, países del sur de Europa como España, Grecia, Italia o Portugal, con un lento crecimiento, siguen padeciendo sus duras consecuencias (Eurostat, 2017). En España, una de las secuelas heredadas de la debacle económica es un mercado laboral tremendamente erosionado. La desigualdad, la pobreza laboral, la temporalidad y el desempleo, sobre todo juvenil, ostentan uno de los porcentajes más elevados de toda la Unión Europea (UE), lo que sitúa a España a la cabeza de la precariedad laboral en Europa (Comisión Europea, 2017; Organización para la Cooperación y el Desarrollo Económicos [OCDE], 2017). Como consecuencia, los trabajadores que manifiestan sentir *siempre* estrés laboral ha aumentado en los últimos 5 años, pasando de una prevalencia del 23% (Eurofound, 2012) a un 30% (Eurofound 2017).

A día de hoy existen infinidad de evidencias científicas que alertan sobre las consecuencias de la exposición a factores de riesgo psicosocial (estresores). Desafortunadamente, la lista de trastornos asociados no está cerrada y recoge múltiples perturbaciones y dolencias. A modo de resumen, de acuerdo con la literatura científica los factores de riesgo psicosocial pueden (1) afectar a la salud mental, manifestándose en forma de trastornos psicológicos, depresión y ansiedad (P. Anderson, Whitmer, Goetzl, Ozminkowski y Wasserman, 2000; Barling y Griffiths, 2003; Bennett, Cook y Pelletier, 2003; Blackmore et al., 2007; Bonde, 2008; Melchior et al., 2007); (2) deteriorar la salud física, manifestándose en forma de enfermedades cardiovasculares (ECV) y gastrointestinales, síntomas y trastornos musculoesqueléticos, problemas de sueño, dolores de cabeza e infecciones (Abekura, et al., 2011; Del Río, 2006; González-Roma et al., 2005; Kivimaki et al., 2002; Monk y Folkard, 1985; Nater et al., 2007); y (3) empobrecer el bienestar laboral, con síntomas tales como insatisfacción laboral,

absentismo, descenso de productividad, aumento de accidentes laborales y una mala calidad del trabajo (Bardera, Osca y González-Camino, 2002; Hart y Wickens, 1990; Lauridsen y Tonnesen, 1990; J.M. Molina, 2010; Salas y Meliá, 2004; Stager, Hameluck y Jubis, 1989). Muchas de estas consecuencias son tan dañinas que empieza a utilizarse el calificativo de *silent killer* (asesino silencioso) para referirse al estrés (Usha, 2016).

De toda la sintomatología asociada al estrés laboral sobresalen, por su prevalencia e incidencia, la depresión y las enfermedades cardiovasculares (ECV). La dimensionalidad del riesgo es tal que la Organización Internacional del Trabajo (OIT, 2012) cifra en más de 150 millones las personas que padecen depresión en el mundo por motivos relacionados con el trabajo, y los pronósticos de la Organización Mundial de la Salud (OMS) auguran que la depresión será la segunda causa de incapacidad laboral a nivel mundial en el año 2020 tras las ECV, que a día de hoy ocupan el primer escalafón con la dramática cifra de 17,5 millones de muertes al año, lo que representa el 31% del total de fallecimientos (OIT, 2016; OMS, 2010).

A pesar de la gravedad del tema que nos ocupa, tan solo un 26% de las empresas europeas dicen haber tomado medidas para reducir el estrés según la Encuesta Nacional de Gestión de Riesgos Laborales en las Empresas (ESENER, 2010). En España, aunque la Ley 31/1995 de Prevención de Riesgos Laborales (LPRL) obliga a eliminar o minimizar todo tipo de riesgos que puedan afectar a la salud de los trabajadores, lo cierto es que se emplean más esfuerzos y recursos para gestionar adecuadamente los factores de riesgo físicos que los de origen psicosocial (García y Olarte, 2010; Jiménez, S, 2014; Merino, 2013). De seguir así, las estadísticas sobre trabajadores afectados “siempre” por estrés (30% según Eurofound, 2017), serán difíciles de disminuir.

Para hacer frente al problema del estrés laboral es necesario identificar y valorar los riesgos de origen psicosocial. Si éstos no se detectan será difícil, por no decir imposible, tomar medidas convenientes para erradicarlos o para minimizar su gravedad; de ahí la importancia de las herramientas de evaluación de riesgos psicosociales como medida preventiva (Moncada y Llorens, 2006; C. Nogareda y Almodóvar, 2006; Oncins y Almodóvar, 1997). Sin embargo, aunque la solución parezca sencilla, abordar este tipo de riesgos no es tarea fácil. De hecho, un 42% de los directivos de la Unión Europea consideran que es más difícil gestionar los riesgos psicosociales que otras cuestiones de seguridad y salud (ESENER, 2010).

Ante estos datos cabe preguntarse por qué un porcentaje tan alto de responsables en prevención consideran espinosa la gestión de los riesgos psicosociales. Hallar la respuesta a esta cuestión, sacando a la luz las principales trabas que la dificultan, será el eje central de la presente tesis. No en vano, Niño (2006) se refiere a la evaluación de estos riesgos como “la ceremonia de la confusión” (p. 30). Como si de una madeja se tratase, se pretende tirar del hilo e ir poniendo nombre a las principales amenazas que complican las evaluaciones de riesgo psicosocial. Entre ellas: (a) la escasa cultura preventiva, revelada por un gran número de organizaciones que aseguran que el principal motivo para valorar los riesgos psicosociales no es prevenir los daños a la salud de los trabajadores, sino evitar multas y sanciones (ESENER-2, 2015); (b) la subestimación de los factores de riesgo psicosocial, lo que propicia que se fije el centro de atención en otro tipo de riesgos como los físicos (Cifre, Salanova y Franco, 2011; Leka y Jain, 2010); (c) la convivencia de distintos intereses y competencias bajo el mismo techo de *prevención*, favoreciendo que se evalúen factores distintos a las condiciones de trabajo, como

la salud, el clima o la satisfacción laboral (Inspección de Trabajo y Seguridad Social [ITSS], 2012; Niño, 2006); o (d) la insuficiente detección y control de posibles fuentes error en el método de evaluación seleccionado (Dupuis et al., 2015), poniendo en tela de juicio la validez de los resultados obtenidos.

A estas circunstancias debe añadirse que, a diferencia de otros riesgos, no existen directrices precisas en forma de guía técnica del Instituto Nacional de Higiene en el Trabajo (INSHT) o normas UNE acerca del modo idóneo de realizar una evaluación psicosocial. Por ello, suelen aplicarse las disposiciones de los artículos 5.3.d) y 5.2 del Real Decreto 39/1997 por el que se aprueba el Reglamento de los Servicios de Prevención (RSP), que establecen que ante esta situación se deben utilizar guías de entidades de reconocido prestigio en la materia u otros métodos o criterios profesionales descritos documentalmente, siempre que proporcionen confianza en el resultado (ITSS, 2012); esto es, que ofrezcan garantías sobre el cumplimiento de ciertos estándares de calidad, como la fundamentación teórica, la validez, la fiabilidad y la baremación.

Como resultado, en las últimas décadas se ha incrementado la confianza en el uso de cuestionarios o autoinformes como técnica principal de evaluación de riesgos psicosociales, ya que, además de ser fáciles de aplicar, de interpretar y de tener un bajo coste (Caicoya, 2004; Ferrer, 2010, ITSS, 2012), están construidas con garantías psicométricas. No obstante, y a pesar de ser un requisito ineludible, no todas las herramientas de evaluación existentes gozan de la misma calidad psicométrica. Ésta es una circunstancia que, por su significación, no deberían pasar por alto las organizaciones y los profesionales implicados en la prevención. De la validez de los resultados va a depender el mayor o menor acierto de las medidas correctoras que se decida llevar a cabo. No hay que olvidar que las decisiones erróneas aquí tomadas repercuten negativamente, tanto en la seguridad y el bienestar de los trabajadores como en la gestión económica de la organización de la que se trate. Conscientes de esta realidad, en esta tesis se ofrecerá información detallada sobre el rigor científico de las técnicas de evaluación de riesgos psicosociales más populares a nivel nacional. El conocimiento de esta información podría simplificar la toma de decisiones a los responsables que deban seleccionar la herramienta de evaluación idónea para su organización.

Con todo, las propiedades psicométricas de fiabilidad y validez no son las únicas cuestiones que deben de tenerse en cuenta, ya que además los test han de carecer de sesgo (Ponsoda, 2009), entendiendo como sesgo “aquellos errores que modifican la información, alejándola de la realidad, y que son inherentes al instrumento usado en la observación o a la metodología usada en su aplicación” (Choi, Granero y Pak, 2010, p.7). Bajo estas premisas, una de las advertencias más reiteradas por los expertos radica en la posibilidad de falseamiento y distorsión de las respuestas de los encuestados, por ser estas medidas sensibles a diversos sesgos, como por ejemplo la aquiescencia, la deseabilidad social o la disimulación, entre otros (De las Cuevas y González de Rivera, 1992; González y Iruarrizaga, 2005; González-Llaneza, 2007; Jensen, 1995). De hecho, la propia Fundación Europea (EuroFound, 2015) desaconseja su uso en las evaluaciones de violencia y acoso laboral o sexual.

Los sesgos de respuesta son un tipo de fuente de error en los test de autoinforme que se refieren a la tendencia a responder un test independientemente del contenido y de la puntuación verdadera del informante (Peer y Gamliel, 2011). Pueden actuar afectando a la fiabilidad y a la

validez de las herramientas de medida, y por consiguiente a la interpretación e inferencias de los resultados (Cheung y Chan, 2002; Cheung y Rensvold, 1999). El sesgo de respuesta que más literatura científica ha generado es la deseabilidad social (DS) (Cosentino y Castro, 2008). Éste es considerado como el sesgo de mayor amenaza en los estudios sobre Psicología y Ciencias Sociales (King y Bruner, 2000). Dicho sesgo se refiere a dar respuestas socialmente aceptables (Crowne y Marlowe, 1960). Por ejemplo, ante la pregunta “*Alguna vez he dicho algo malo de otra persona*”, la respuesta negativa sería la que proporcionaría una imagen positiva ante la sociedad. Así, dependiendo de la impresión que se quiera ofrecer, la distorsión puede ser positiva o negativa. Aunque existe controversia acerca de si la DS es un sesgo que incrementa el error de medida (Crowne y Marlowe, 1964; Furnham, 1986; Paulhus, 1991) o de si es un rasgo de personalidad (Ferrando y Chico, 2000; Fleming y Zizzo, 2011; Paulhus y John, 1998; Pauls y Stemmler, 2003), de lo que no cabe duda es que su efecto distorsiona las mediciones (Morales, 2006) y que su control ayudaría a minimizar la varianza de error (Anguiano-Carrasco, Vigil-Colet y Ferrando, 2013).

Los sesgos de respuesta no son los únicos *enemigos* de las medidas autoinformadas. Dentro de las amenazas a la validez también ha entrado en juego el afecto negativo (AN), un estado emocional capaz de influir en las puntuaciones de los test. Se trata de una dimensión bipolar donde los niveles bajos representan estados de calma y serenidad, en contraposición a los niveles altos que describen estados de ira, culpa, temor y nerviosismo (Watson, Clark y Tellegen, 1988). Los sujetos con altos niveles de AN reportan una visión negativa del mundo que los rodea y de ellos mismos, experimentando mayores amenazas (J. Martín, Luceño, Jaén y Rubio, 2007). De ese modo, perciben exposiciones de estrés más periódicas y de mayor grado (Sandi y Calés, 2000), de ahí que el AN haya explicado la varianza encontrada en cuestionarios de medidas estrés, estresores o síntomas psicósomáticos, entre otros (Brief, Burke, George, Robinson y Webster, 1988; Watson y Pennebaker, 1989). La propia EU-OSHA (2005) también ha señalado el AN como variable de confusión que influye en la percepción de las condiciones laborales.

Ningún test está exento de fuentes de error (F. Jiménez y Sánchez, 2004), sin embargo, que contenga ítems sesgados no implica que toda la escala en su conjunto esté sesgada (Delgado, 1998), pues existen medidas para detectar y equilibrar esos reactivos (Barker, Pistrang y Elliott, 2002; Echeburúa, Amor y Corral, 2003; Morales, 2006). Toda una rama científica, la Psicometría, está dedicada a desarrollar técnicas de detección y minimización de la varianza originada por los sesgos. Desde las escalas explícitas de detección de sesgos hasta procedimientos estadísticos de tratamiento de datos, las alternativas de control de sesgos son muy variadas (véase Morales, 2006 para un conocimiento más específico). Sin embargo, son pocas las investigaciones que controlan sus efectos (Dupuis et al., 2015) y, de ellas, la mayor parte se focalizan en evaluaciones de personalidad.

Estos y otros obstáculos serán expuestos sobre la mesa para su estudio y constituirán el foco principal de esta tesis. Para ello, se examinará pormenorizadamente la literatura científica en busca de las principales amenazas a la validez de las medidas autoinformadas, haciendo especial hincapié en la disponibilidad y en la calidad de los instrumentos de evaluación de riesgos psicosociales, pues se cree que aquí reside gran parte del dilema (Peiró, 1999a). Algunos de los motivos son descritos por Meliá (2008): (1) lo novedoso de estos riesgos; (2) la escasez

de formación especializada por parte de muchos profesionales dedicados a la prevención, en disciplinas tan necesarias como la evaluación psicológica, psicometría, psicología del trabajo o de la seguridad y salud ocupacional; y (3) la celeridad con la que algunas veces se busca una respuesta práctica, o lo que es peor, interesada. Según el citado autor, este conjunto de circunstancias ha provocado “procesos de evaluación poco fundados o el uso poco fundado de instrumentos y métodos con fundamentos razonables” (párr. 4). Al respecto, se examinarán las herramientas de evaluación más populares a nivel nacional.

Los hallazgos derivados de la revisión literaria serán ordenados a modo de proposiciones que, mediante la aplicación de un riguroso método Delphi, serán contrastadas por un elenco de expertos y profesionales en la materia, conformando así el primero de los estudios de esta tesis. Fruto del juicio y el consenso realizado por los expertos, podrán identificarse las principales fuentes que entorpecen y afectan a la validez de las evaluaciones de riesgo psicosocial, permitiendo realizar una composición de lugar sobre el estado actual de dichas evaluaciones. Al mismo tiempo, los participantes en el estudio brindarán recomendaciones para la mejora de las mismas.

El logro de estas metas beneficiaría a una amplia mayoría de profesionales dedicados a evaluar riesgos psicosociales, debido a su gran utilidad práctica. No obstante, se pretende desgranar todavía más los resultados y explorar parte de los focos de controversia destapados por los miembros del panel Delphi, mediante un segundo estudio de corte cuantitativo. Concretamente, se analizará el impacto de la deseabilidad y el afecto negativo sobre los cuestionarios de evaluación de riesgo psicosocial. Estas cuestiones no están debidamente investigadas por la comunidad científica, manifestación de lo cual es la escasa literatura disponible al respecto. Como consecuencia, no hay suficientes evidencias científicas que permitan conocer o contrastar la influencia de los sesgos de respuesta en las medidas de evaluación de riesgos psicosociales, ni tampoco es posible asegurar si el control de éstos mejoraría la calidad de los resultados. Ahondar en este reto constituirá el segundo de los retos de esta tesis.

En síntesis, y a modo de resumen, los objetivos planteados pueden concretarse en dos:

Objetivo 1: Explorar los motivos que dificultan las evaluaciones de riesgos psicosocial.

- a) Desde el punto de vista experto, elaborar una composición de lugar de la problemática que envuelve las evaluaciones de riesgos psicosociales en el trabajo.
- b) Explorar las debilidades y fortalezas de dichas evaluaciones cuando son medidas con autoinformes.
- c) Extraer directrices prácticas que consigan minimizar los puntos flacos de las evaluaciones psicosociales y que, a su vez, puedan servir de orientación para el ejercicio profesional.

Objetivo 2: Comprobar la existencia e influencia de los sesgos de respuesta en las medidas de los riesgos psicosociales en el trabajo.

- a) Determinar si la DS tiene algún efecto en dichas medidas.
- b) Analizar el funcionamiento diferencial del ítem para analizar posibles sesgos producidos por el rasgo de AN.

La estructura de este trabajo de investigación se ha modelado en seis capítulos. El primero presenta una aproximación conceptual sobre los factores de riesgo psicosocial. Además, se realiza una descripción detallada de los principales riesgos psicosociales y sus consecuencias tanto para la salud de los trabajadores como para la propia organización, en un intento de subrayar la gran importancia que implica evaluar los riesgos psicosociales como herramienta preventiva capaz de evitar la materialización de dichos daños. Se finaliza el capítulo con una reflexión sobre el panorama actual y las perspectivas de futuro de los riesgos y los factores psicosociales.

El segundo capítulo se centra en la evaluación de los riesgos psicosociales, incluyendo una revisión exhaustiva de la literatura científica acerca de la normativa aplicable y las recomendaciones para su medición, especificando a su vez los requisitos que necesita cumplir toda herramienta de evaluación para ofrecer garantías psicométricas. En esa dirección se incluye una revisión de los principales instrumentos de medida utilizados en el marco preventivo, estableciendo sus principales ventajas y desventajas.

El tercer y último capítulo del marco teórico profundiza en las amenazas a la validez de las medidas autoinformadas desde el enfoque de los sesgos de respuesta y de otras variables de confusión, que puedan estar actuando como distorsionadoras de la realidad laboral alterando los resultados de las evaluaciones de riesgos psicosociales.

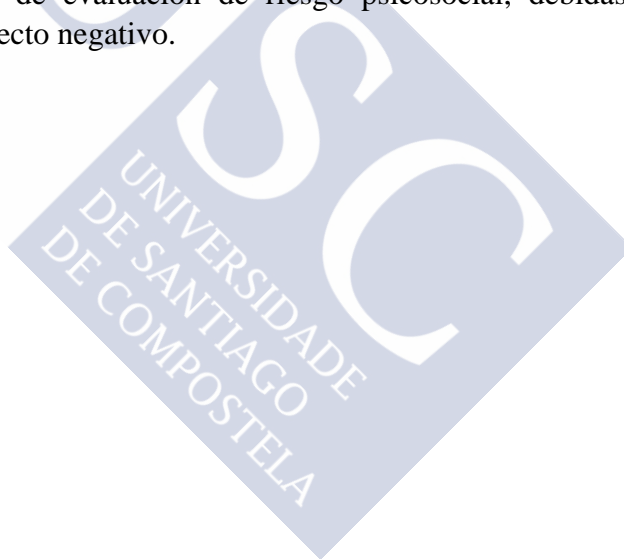
En conjunto, los tres primeros capítulos configuran el marco conceptual que fundamenta los posteriores estudios empíricos de la tesis. Así, el capítulo IV intenta contrastar las principales reflexiones derivadas de la revisión de la literatura para arrojar algo de luz sobre cuáles son los principales obstáculos que están entorpeciendo las evaluaciones de riesgos psicosociales. La estrategia seleccionada para ello consistió en la aplicación de la técnica mixta (cualitativa-cuantitativa) del panel Delphi, donde se somete al juicio experto determinadas conclusiones referentes a la evaluación de riesgos psicosociales.

Como consecuencia de las valoraciones y aportaciones realizadas por los expertos en el Delphi, se creó el segundo de los estudios (capítulo V). De tipo exploratorio y de corte cuantitativo, su objetivo es profundizar en las cuestiones que los participantes expertos en el Delphi señalan como las más controvertidas. En concreto, trata de averiguar qué papel juegan la deseabilidad y el afecto negativo como potenciales distorsionadores de las puntuaciones de los cuestionarios que se utilizan para medir riesgos psicosociales. Para comprobar la influencia de estos sesgos se utilizarán técnicas avanzadas de análisis factorial (Ferrando, 2005) y salvar, a su vez, las limitaciones de determinados métodos de control, como por ejemplo las escalas específicas de detección de sesgos, cuya crítica principal se centra en el método de corrección empleado que elimina las puntuaciones de los sujetos con altos niveles en la variable sin garantizar que no son puntuaciones verdaderas (Morales, 2006). Para alcanzar estos objetivos se partirá de una muestra multiocupacional de trabajadores a nivel nacional. La estrategia consiste en tratar los datos estadísticamente para detectar y eliminar sesgos de respuesta como la deseabilidad social, siguiendo el método basado en el análisis factorial propuesto por Ferrando, Lorenzo-Seva y Chico (2009). Esta técnica permite obtener puntuaciones limpias de sesgo sin eliminar información de interés. A su vez, y como segunda parte del estudio, se analizará la influencia del AN en las puntuaciones del cuestionario de evaluación de riesgos psicosociales, comprobando el funcionamiento diferencial de los ítems y utilizando como grupo

focal los sujetos con niveles altos de afecto negativo. Lo que se pretende averiguar es si existen ítems en el test donde la posibilidad de respuesta sea distinta, a favor o en contra, en función de la pertenencia al grupo de afecto negativo.

De cada estudio se exponen los objetivos principales y la metodología, se describen exhaustivamente los resultados obtenidos y las conclusiones, y finalmente se comentan las limitaciones encontradas, planteando recomendaciones prácticas, así como directrices para la mejora de futuros trabajos en este campo. El último de los capítulos (capítulo VI) recogerá una síntesis de las conclusiones generales extraídas de toda la tesis en su conjunto.

Dada la escasez de estudios, la necesidad de ampliar conocimientos y arrojar luz en este campo se hace primordial. A ello aspira la presente tesis, que se plantea como una propuesta de mejora de la que pueden emerger nuevos conceptos a explorar y, poco a poco, ir allanando el camino a futuras investigaciones que, desde el marco preventivo, se aventuren a facilitar a los profesionales del sector un mejor entendimiento del complejo fenómeno de la evaluación psicosocial laboral. Los hallazgos derivados de la consecución de estos retos permitirán a los profesionales e investigadores en el campo de la prevención de riesgos utilizar orientaciones prácticas basadas en la evidencia para mejorar la validez de las evaluaciones de riesgo psicosocial, y también cuantificar el grado de distorsión de respuesta de los trabajadores en las puntuaciones de los test de evaluación de riesgo psicosocial, debidas al impacto de la discapacidad social y al afecto negativo.





PARTE 1: MARCO TEORICO





1. FACTORES DE RIESGO PSICOSOCIAL vs RIESGOS PSICOSOCIALES EN EL TRABAJO

1.1. FACTORES PSICOSOCIALES DE RIESGO.

Los factores psicosociales no son fáciles de entender, son cuestiones complejas basadas en la percepción subjetiva de los trabajadores, que a su vez contemplan muchos otros aspectos (OIT, 1986). Por ende, la identificación del término *factor psicosocial* tampoco está exenta de dificultad. Sirva de muestra la confusión existente a la hora de distinguir entre factor psicosocial, factores psicosociales de riesgo y riesgo psicosocial. En un intento por arrojar un poco de claridad sobre este desconcierto, se abordará cada uno de estos términos. Para comenzar, a continuación se recogen las definiciones más citadas sobre el concepto “factor psicosocial”.

Se ha seleccionado en primer lugar la cita del comité Mixto (OIT)-(OMS), que en su novena reunión (1984) delimita los factores psicosociales como:

Interacciones entre el trabajo, su medio ambiente, la satisfacción en el trabajo y las condiciones de su organización, por una parte, y por la otra, las capacidades del trabajador, sus necesidades, su cultura y su situación personal fuera del trabajo, todo lo cual, a través de percepciones y experiencias, pueden influir en la salud y en el rendimiento y la satisfacción en el trabajo (p.3).

Otra definición muy utilizada es la que se recoge en la Nota Técnica de Prevención nº 443 del INSHT (Martín-Daza y Pérez, 1997), donde estos factores se describen de la siguiente manera:

Aquellas condiciones que se encuentran presentes en una situación laboral y que están directamente relacionados con la organización, el contenido del trabajo y la realización de la tarea, y que tienen capacidad para afectar tanto al bienestar o a la salud (física, psíquica o social) del trabajador como al desarrollo del trabajo (p.1).

Para Lahera y Góngora (2002) significan:

Aquellas condiciones que se muestran presentes en una situación laboral y que están directamente relacionadas con la organización, el contenido del trabajo y la realización de la tarea, y que poseen capacidad para afectar tanto al desarrollo del trabajo como a la salud física, psíquica o social del trabajador (p.5).

Las definiciones que abordan el significado de factor psicosocial son muy numerosas (Carayon, Haims y Yanf, 2001; Cox y Griffiths, 1996; González-Trijueque, Giachero y Delgado, 2012; Kalimo, 1988; Moncada y Artaco, 2000), y desentrañarlas una a una carece de

sentido cuando todas coinciden al concluir que se trata de condiciones derivadas de la organización del trabajo con posibilidad de afectar a la salud y al bienestar de los trabajadores.

Es preciso aclarar que cuando se habla de influir en la salud, no hay que interpretarlo siempre como una intervención negativa, de hecho, los factores psicosociales pueden actuar sobre la salud de los trabajadores desde una doble perspectiva. Primeramente, se encuentra el enfoque positivo de mejora de la salud que está determinado por la existencia de un equilibrio entre (1) las capacidades y necesidades humanas; y (2) las condiciones de trabajo, la motivación y la satisfacción laboral (Gil-Monte, 2009; OIT, 1984). Por ejemplo, un buen clima de trabajo o un buen jefe pueden facilitar el desarrollo personal y profesional del trabajador, aumentando su satisfacción laboral (Kalimo, 1988). La segunda vertiente alude a un efecto negativo que se origina cuando las condiciones de la organización y gestión del trabajo constituyen un factor de riesgo para la salud del trabajador, al generar respuestas fisiológicas de tensión y estrés con alta probabilidad de causar daños físicos, psíquicos o sociales (Benavides et al., 2002; Bucci, 2012; Cox y Griffiths, 1996; Cox, Griffiths y Leka, 2005; Gil-Monte, 2014; Moreno y Báez, 2010).

Por lo tanto, y parafraseando la conceptualización de Moreno (2011, p.7), los factores psicosociales en el trabajo son meros “descriptivos” que se refieren a las condiciones laborales de la empresa (ej., clima laboral, estilo de mando, diseño del puesto de trabajo) pudiendo ser positivos o negativos. Cuando son desfavorables, actúan como “predictivos” al llevar implícita la posibilidad de dañar la salud y el bienestar de los trabajadores. En este último caso son generadores de tensión o, dicho de otro modo, estresores organizacionales que comúnmente se alude a ellos con el término de *factores de riesgo psicosocial* (Benavides et al., 2002; Carrasco, de la Corte y León, 2010; Gil-Monte, 2010; Moreno, 2011).

Precisamente, por su carácter dañino, los factores psicosociales adversos (estresores) han sido los que más investigados, en comparación con los factores de influencia positiva. No obstante, existen variables individuales que juegan un papel importante en la aparición de estrés, entre ellas la introversión, la autoestima o la inestabilidad emocional. Son los llamados estresores individuales. Por lo tanto, los resultados desfavorables van a depender de las características personales y de la capacidad del trabajador para enfrentarse a los desajustes causados (NTP 318 de Martín-Daza, 1993). Asimismo, puede ocurrir que ante una misma exposición a un factor psicosocial adverso la respuesta de los trabajadores puede ser diferente; es decir, mientras en unos casos tendrá consecuencias perjudiciales para la salud, en otros puede no haber repercusiones (NTP 443 de Martín-Daza y Pérez-Bilbao, 1997; OIT, 1984).

A la hora de establecer una clasificación de los factores de riesgo psicosocial, pueden consultarse gran variedad de propuestas, como las expuestas por Cooper (1993), Leka, Griffiths y Cox (2004), O’Brien (1986) o Peiró (2005). Al ser múltiples las causas que pueden incidir en su origen se trata de una lista abierta, lo que convierte en inverosímil el intento de abarcarlos todos. En este sentido, y a modo de síntesis, se muestra en la Tabla 1 una categorización no exhaustiva de los más citados.

Tabla 1. Factores de Riesgo Psicosocial en el Trabajo
Fuente: Adaptado de Mansilla y Favieres (2011) y Sauter, Hurrell, Murphy y Leví (2001).

FACTORES DE RIESGO PSICOSOCIAL EN EL TRABAJO		
Entorno laboral	Organización del tiempo de trabajo	Contenido, gestión y organización del trabajo
Agentes físicos y biológicos	Horario de trabajo	Monotonía y repetitividad de la tarea
Iluminación	Trabajo a turnos	Sentido del trabajo
Ruido ambiental	Nocturnidad	Ritmo de trabajo
Vibraciones	Pausas y descansos	Conflicto y Ambigüedad de rol
Temperatura	Flexibilidad horaria	Autonomía
Radiaciones		Cualificación profesional
Virus		Carga mental
Bacterias		Formación
Hongos		Desarrollo profesional
Diseño del puesto de trabajo		Cultura organizativa y forma de funcionar
Adaptación ergonómica		Estilo de mando
Configuración espacial		Participación en la toma de decisiones
		Relaciones interpersonales en el trabajo
		Condiciones de empleo
		Responsabilidad
		Relación trabajo-casa

Tal y como se desprende de la tabla 1, todos los factores recogidos están directamente relacionados con la organización y gestión del trabajo, y es su posibilidad de generar estrés e influir negativamente en la salud del trabajador lo que los convierte en factores de riesgo psicosocial, dicho de otro modo, son peligros psicosociales. A lo largo de este primer capítulo se hará referencia explícita a muchas de las evidencias científicas que han demostrado la capacidad de daño de los estresores.

Hasta aquí, se ha intentado clarificar la distinción entre factor psicosocial y factor de riesgo psicosocial, detallando ambos conceptos. Creyendo haber cumplido el propósito, toca hacer mención explícita del término de *riesgo psicosocial* (de origen laboral).

1.2. RIESGO PSICOSOCIAL

Antes de profundizar en la descripción de *riesgo psicosocial laboral*, es preciso detenerse ante el concepto de *riesgo laboral*, el cual es definido por ley (art. 4, LPRL) como “la posibilidad de que un trabajador sufra un determinado daño derivado del trabajo”. Si se presta debida atención al artículo 7.2 de la misma ley, donde se especifica que las *condiciones de trabajo* son “todas aquellas otras características del trabajo, incluidas las relativas a su organización y ordenación, que influyan en la magnitud de los riesgos a los que esté expuesto el trabajador”, se incluye explícitamente la esfera psicosocial dentro de la noción de riesgo laboral.

En consecuencia, el riesgo psicosocial derivado del trabajo es, sin lugar a dudas, un riesgo laboral. Llana (2009) lo define como la posibilidad de que un trabajador sufra un daño derivado de los factores psicosociales que, como se ha señalado anteriormente, se encuentran

en las características de la organización (ej., estilo de mando, carga de trabajo, autonomía). Un suelo mojado y resbaladizo es un peligro observable que proporciona una idea clara, tanto de la probabilidad de caída como de su capacidad de daño, por lo que es un riesgo fácilmente evaluable. Sin embargo, la naturaleza a menudo subjetiva, invisible y poco observable de los peligros psicosociales (estresores), así como de sus efectos sobre la salud, convierten la evaluación del riesgo en una tarea compleja. Este reto de medir los riesgos de origen psicosocial será abordado en el siguiente capítulo de este proyecto.

En la tabla 1 se han enunciado algunos de los peligros psicosociales que pueden estar presentes en una organización; como todo peligro lleva asociado un riesgo, es menester aludir a los riesgos psicosociales más citados, y con ese objeto se elabora la tabla 2.

Tabla 2: Riesgos Psicosociales
(Fuente: Elaboración propia)

RIESGOS PSICOSOCIALES
Estrés laboral
Desgaste profesional o Burnout
Violencia física y psíquica
Conciliación laboral-familiar
Acoso laboral (Mobbing) y Acoso Sexual
Inseguridad laboral
Trabajo emocional (Autocontrol de las emociones)

El Instituto Sindical de Ambiente Trabajo y Salud (ISTAS, 2010) expresa muy bien la diferencia entre conceptos:

En términos de prevención de riesgos laborales los factores psicosociales representan la exposición (o sea: lo que habrá que identificar, localizar y medir en la evaluación de riesgos), la organización del trabajo el origen de ésta (o sea: sobre lo que habrá que actuar para eliminar, reducir o controlar estas exposiciones), y el estrés, el precursor o antecesor del efecto (enfermedad o trastorno de salud) que se pretende y debe evitar (p.18).

De lo expuesto, se deduce que los factores de riesgo psicosocial y los riesgos psicosociales son nociones distintas. Para concluir con la diferenciación entre ambos, podemos afirmar que los factores de riesgo psicosocial son peligros psicosociales (Cox y Griffiths, 1996; Cox, Griffiths. y Randall, 2003), ya que constituyen una fuente de daño potencial cuyo origen es de tipo organizativo, como la exposición a ritmos de trabajo excesivo, trato injusto, futuro inseguro, falta de apoyo social o falta de control, entre otros factores, con distintos niveles de probabilidad de afectar la salud de los trabajadores. Sin embargo, los riesgos psicosociales son hechos (respuestas del organismo), causados por las condiciones de la organización, con alta posibilidad de alterar y dañar gravemente la salud de los trabajadores (Moreno y Báez, 2010). Por lo tanto, los riesgos psicosociales de origen laboral son situaciones originadas por la exposición a factores psicosociales adversos directamente relacionados con la organización del trabajo. Dicho en otras palabras, los factores psicosociales de riesgo son antecedentes de los

riesgos psicosociales (Luceño, Martín, Rubio y Díaz, 2010; Llaneza, 2009; Moncada et al., 2011; Moreno y Báez, 2010; Peiró, 2008).

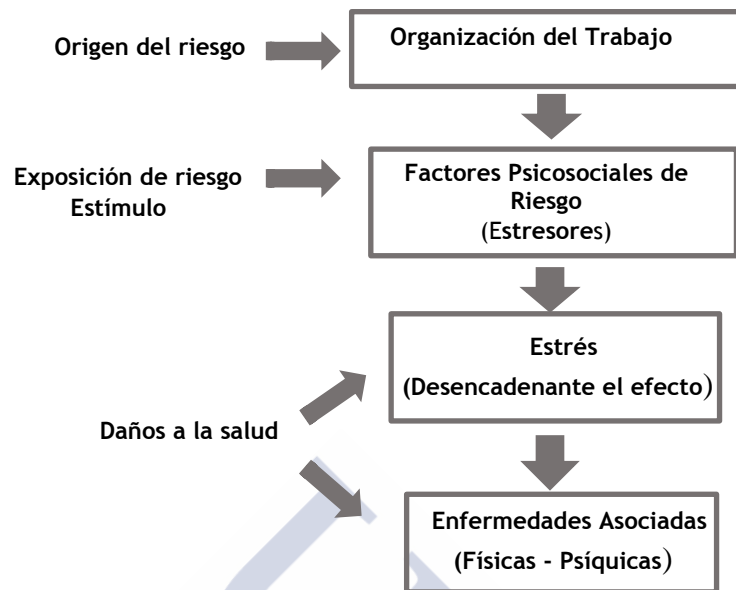
1.3. LOS FACTORES PSICOSOCIALES Y SU RELACIÓN CON LA SALUD

De acuerdo con el Preámbulo de la Constitución de la OMS (1948, p. 1), “la salud es un estado de completo bienestar físico, mental y social, y no solamente la ausencia de afecciones o enfermedades”. Dentro de esta consideración positiva de la salud, entendida desde una perspectiva global y no sólo como ausencia de daño, cabría mencionar que un ambiente de trabajo saludable no es sólo aquel que carece de factores perjudiciales para la salud, sino también aquél que contiene elementos capaces de promocionar la salud (Leka et al., 2004). A la luz de este marco, se encuadran los estudios sobre factores psicosociales con efectos positivos sobre el bienestar y la salud del trabajador. A pesar de ser aún escasas, en los últimos años han repuntado las investigaciones que desarrollan nuevos constructos derivados de la Psicología positiva, como el *Engagement*, cuya formulación teórica sería contraria al *burnout* (Carrasco et al., 2010, Pocinho y Perestrelo, 2011; Salanova y Schaufeli, 2009; Salanova, Schaufeli, Llorens, Peiró y Grau, 2000; Ten Brummelhuis, Bakker, Hetland y Keulemans, 2012). Estos factores de influencia positiva no son analizados por las evaluaciones de riesgo psicosocial amparadas por la ley 31/1995 (LPRL), por ese motivo en esta tesis no serán objeto de análisis.

Lo que tradicionalmente ha predominado es el enfoque “negativo” de la salud laboral (Díaz, Blanco, Horcajo y Valle, 2007), es decir, el estudio de cómo los factores de riesgo psicosocial -a saber, tareas monótonas o triviales, trabajos con plazos estrictos, horario de trabajo imprevisible, falta de control, no participar en la toma de decisiones, inseguridad laboral o malas relaciones con los compañeros, entre otros factores (Luceño, Martín, Rubio, et al., 2008; OMS, 2004)-, ejercen una influencia negativa sobre el bienestar de los trabajadores alterando su salud.

La literatura revisada, con mayor apoyo empírico y teórico, ha revelado que el riesgo de tensión laboral aumenta en los lugares de trabajo donde (1) confluyen el exceso de exigencias psicológicas, la falta de control y escaso apoyo social por parte de compañeros y superiores (Karasek, 1979, Karasek y Theorell, 1990); (2) se da un desequilibrio entre esfuerzo y recompensas (Siegrist, 1998); o (3) existe falta de confianza o justicia organizacional (Moncada, S., Llorens, C., Andrés, R., Moreno, N. y Molinero, 2014), entre otras muchas circunstancias. De este modo, cuando los requerimientos del entorno laboral obligan a realizar un esfuerzo que sobrepasa la capacidad de adaptación del trabajador, se ocasionan una serie de problemas de ajuste que pueden dar lugar a la aparición de estrés. Un esquema que ilustra de manera resumida la relación de estos factores y la salud se muestra en la siguiente figura.

Figura 1: Relación factores de riesgo psicosocial - enfermedad
Fuente: Adaptación de Martínez- Plaza (2007, 2010)



Existen numerosas evidencias científicas que han relacionado determinados factores de riesgo psicosocial con la salud de los trabajadores. Así, en lo que concierne a las altas exigencias psicológicas se ha demostrado que el riesgo para la salud será mayor cuanto menor sea la autonomía y el apoyo social de compañeros y superiores (Moncada, Llorens y Kristensen, 2002). Al respecto, Johnson y Hall (1988, 1994) evidenciaron que estos trabajadores desarrollan un riesgo de mortalidad por enfermedad cardiovascular dos veces mayor que los trabajadores expuestos a bajas exigencias, mucho control y mayor apoyo social. En la misma línea, De Jonge, Bosma, Peter y Siegrist (2000) han relacionado los trastornos de ansiedad, fobias, obsesiones, fatiga y falta de ánimo con el exceso de exigencias psicológicas cuantitativas y la falta de autonomía.

En lo que respecta a la falta de apoyo social, De Beeck y Hermans (2000) indicaron que existe mayor posibilidad de afectar a los trabajadores con baja satisfacción laboral y con elevadas demandas laborales. Se ha relacionado a su vez con el estrés y el tabaquismo (Adriaanse, Van Reek, Zandbelt y Evers, 1991; Mangione y Quinn, 1975). Por otra parte, se ha demostrado que el apoyo de compañeros y superiores atenúa los efectos que las exigencias del trabajo producen sobre el organismo, como la depresión o la insatisfacción laboral (Karasek, Triantis y Chaundhry, 1982) y que el número de horas trabajadas se relaciona positivamente con el apoyo social percibido (Fielden y Peckar 1999).

También han sido objeto de numerosas investigaciones la calidad de liderazgo, la falta de control y la inestabilidad en el empleo. Los estudios relacionados con el primer factor han confirmado que los directivos que muestran atención hacia sus empleados y les permiten una mayor autonomía, ocasionan mayor bienestar y satisfacción laboral en los trabajadores (Cooper y Smith 1986). En lo concerniente a la baja autonomía de decisión o el escaso control sobre el trabajo, son diversos los estudios que han manifestado correlación con estrés, ansiedad, depresión, apatía y agotamiento, baja autoestima y un aumento de trastornos cardiovasculares (Karasek y Theorell, 1990; Sauter, Hurrell y Cooper, 1989). En cuanto a la inestabilidad en el

empleo, ésta se ha ratificado como un riesgo para la salud (Ferrie, 1999; Quinlan y Mayhewhle 2001) asociándolo al colesterol, la hipertensión, o la obesidad (Ferrie, Shipley, Stansfeld, y Marmot, 2002), así como a lesiones laborales y trastornos mentales (Virtanen et al., 2005). En la misma dirección apunta Siegrist (1996a), quien señala que los trabajadores expuestos a esfuerzos elevados, salarios inadecuados y bajo control disminuyen su nivel de autoestima considerablemente.

1.4. PRINCIPALES RIESGOS PSICOSOCIALES EN EL TRABAJO Y SU RELACIÓN CON LA SALUD

Entre los riesgos psicosociales más reconocidos, divulgados y estudiados dentro del ámbito laboral se encuentran el estrés laboral, el síndrome de estar quemado o *burnout*, y el acoso laboral o *mobbing* (González-Trijueque et al., 2012; Moreno y Báez, 2010). Dada su importancia, en los próximos epígrafes se pretende realizar una detallada descripción de cada uno de ellos.

1.4.1. Estrés laboral.

Como se ha visto en líneas precedentes, el estrés es señalado como la consecuencia más directa a la exposición de los factores psicosociales adversos, motivo que hace imprescindible adentrarse en este término.

1.4.1.1. Delimitación conceptual

La definición del término estrés, tal y como se entiende hoy, fue acuñada por el médico checo Hans Selye, quien en 1936 lo definió como la reacción o respuesta no específica del organismo ante cualquier demanda de cambio. Llegó a esta conclusión después de experimentar con numerosos animales de laboratorio, sometidos a distintos estímulos especialmente dañinos, como inyecciones de toxinas, calor o frío. Observó que todas las sustancias tóxicas, cualquiera que fuese su origen, producían la misma respuesta y mostraban cambios fisiológicos muy parecidos a los de los humanos, como ataques cardíacos, artritis, úlceras gástricas o fallos renales, entre otros. Esta sintomatología formaba un síndrome bien definido al que denominó Síndrome General de Adaptación. (SGA, 1936), enmarcando como estrés el estado manifestado por el SGA y como estresor al estímulo desencadenante de esa respuesta.

En 1967 los psiquiatras Thomas Holmes y Richard Rahe investigaron la relación existente entre los eventos estresantes y las causas de enfermar. Examinaron los expedientes médicos de más de 5000 pacientes para determinar si los acontecimientos estresantes provocaban enfermedades, y observaron que años antes del desarrollo de la enfermedad habían sufrido un gran número de diferentes situaciones potencialmente estresantes. A partir de esos datos realizaron un inventario de 43 sucesos vitales estresantes, siendo los resultados publicados como Escala de Rango de Reajuste Social (*Social Readjustment Rating Scale*), también conocida como Escala de Estrés de Holmes y Rahe. Estos investigadores constataron que si en un periodo de dos años la suma a los sucesos vitales estresantes es superior a los 300 puntos el riesgo de desarrollar una enfermedad asociada al estrés sería de un 80%; si la suma se encuentra entre 200 y 299 puntos, supondría un 50% de riesgo, y si está entre 150 y 199 puntos el riesgo sería de un 33%. Para un mejor entendimiento de estos datos, en la Tabla 3 se muestra dicha escala con los 43 eventos potencialmente estresores.

Tabla 3: Escala de estrés de Holmes y Rahe
Fuente: Adaptado de Holmes y Rahe (1967)

Acontecimiento	Valor estrés	Acontecimiento	Valor estrés
Muerte de la pareja	100	Dormir menos de 8 horas	25
Divorcio	60	Cambio de responsabilidades en el trabajo	25
Menopausia	60	Problemas con la familia política o hijos	25
Separación de la pareja	60	Logro personal excepcional	25
Encarcelamiento, pérdida de libertad	60	La pareja comienza o deja de trabajar	20
Muerte de un pariente cercano	60	Comenzar o terminar la escuela	20
Enfermedad o incapacidad	45	Cambios en las condiciones de vida	20
Matrimonio	45	Cambio en hábitos personales	20
Pérdida de empleo	45	Alergia crónica	20
Reconciliación con la pareja	40	Problemas con el jefe	20
Jubilación	40	Cambio en horario o condiciones laborales	15
Cambio de salud en pariente cercano	40	Cambio de casa	15
Trabajar más de 40 horas por semana	35	Síndrome pre-menstrual	15
Embarazo	35	Cambio de escuela	15
Problemas sexuales	35	Cambio de actividad religiosa	15
Llegada de un nuevo miembro a la familia	35	Cambio de actividades sociales	15
Cambio de puesto laboral	35	Préstamo menor	10
Cambio de estado financiero	35	Cambio en frecuencia de reuniones familiares	10
Muerte de un amigo cercano	30	Vacaciones	10
Aumento de las discusiones con la pareja	30	Infracción menor de la ley	10
Hipoteca o préstamo	25	Vacaciones navideñas	10
Problema con hipoteca o préstamo	25		

Nota: Se han resaltado en gris aquellos acontecimientos vitales directamente relacionados con el mundo laboral.

Tal y como se observa, Holmes y Rahe (1967) han incluido en su lista de estresores vitales a las circunstancias directamente relacionadas con el trabajo. Así, ampliando el concepto a la esfera laboral, que es la que nos ocupa, las situaciones estresantes serían los factores psicosociales de riesgo que actuarían como potenciales desencadenantes de estrés laboral; esto es, cuando las exigencias del entorno laboral obligan a realizar un esfuerzo que sobrepasa la capacidad de adaptación, provocando en el organismo una serie de mecanismos de ajuste que pueden causar estrés, actuando éste a su vez como precursor de daños a la salud, tanto física como psíquica (Leka y Jain, 2010; OIT, 2012).

A pesar de existir numerosos estudios acerca del estrés, su conceptualización está repleta de controversia, motivo que le impide gozar de consenso a la hora de definirlo (P. Martín, Salanova y Peiró, 2003). En líneas generales, se ha utilizado el término estrés desde tres aproximaciones: (i) como estímulo (Canon, 1932; Holmes y Rahe, 1967), (ii) como respuesta (Selye, 1960), y (iii) como interacción entre la persona y el entorno (Lazarus y Folkman, 1986), que a su vez es el enfoque más utilizado. Otros investigadores como Cox y Macay (1981) agregan un nuevo punto de vista: *el estrés como percepción*. Entre tantas definiciones la que goza de gran aprobación es la propuesta por McGrath (1970): “el estrés es un desequilibrio sustancial (percibido) entre la demanda y la capacidad de respuesta (del individuo) bajo condiciones en las que el fracaso ante esta demanda posee importantes consecuencias (percibidas)” (p. 20). Siguiendo la misma línea, Leka et al. (2004), definen el estrés laboral como “la reacción que puede tener el individuo ante exigencias y presiones laborales que no se

ajustan a sus conocimientos y capacidades, y que ponen a prueba su capacidad para afrontar la situación” (p. 12).

Por tanto, el estrés sería el conjunto de reacciones que experimenta el organismo como respuesta ante determinados estímulos. En contra de lo que a veces se piensa, tales reacciones no tienen por qué ser siempre negativas. Es erróneo equiparar el estrés con enfermedad, pues éste puede contener también connotaciones positivas. En ese caso estaríamos hablando del denominado eustrés o estrés positivo, que actúa como mecanismo de defensa protegiendo el organismo (Hamberger y Lohr, 1984; Selye, 1974).

No obstante, el objeto de estudio de esta tesis solamente lleva a considerar dentro del estrés laboral, *el distrés* o estrés con consecuencias negativas. Esto es debido a que se trata de un riesgo laboral que para hacer frente a las exigencias del trabajo provocaría en el organismo reacciones fisiológicas, psíquicas y de comportamiento desfavorables para la salud. (L. Sánchez, Simoes y Brito., 2008). Como todo riesgo laboral, su evaluación es obligada por ley (LPRL) con la finalidad de eliminarlo o minimizarlo, y en ese sentido las evaluaciones de riesgos psicosociales serán el fundamento del presente trabajo de investigación.

1.4.1.2.El estrés laboral como precursor de enfermedad

Cuando una persona percibe una amenaza, tanto real como imaginaria, el sistema nervioso central envía señales de alarma al organismo para prepararse. Es en ese momento cuando se activan distintos mecanismos funcionales para facilitar la huida, el ataque o la experiencia de miedo, como por ejemplo el aumento de la presión sanguínea, de la frecuencia respiratoria, del ritmo cardíaco y de la insulina, además de la liberación de nutrientes, de mayor oxigenación muscular, del incremento de la producción de plaquetas o de la desviación de la sangre hacia los músculos (OIT, 2012). Este proceso ha sido el responsable de la supervivencia del ser humano ante las graves amenazas. Por ejemplo, sin él nuestros antepasados no hubiesen podido evitar ser devorados ante el encuentro con un león en la sabana africana, debido a que la falta de estrés no permite activar los mecanismos necesarios en el organismo para facilitar la huida o la lucha (Yanes, 2008; Sapolsky, 2004). Esta concatenación de reacciones fisiológicas está preparada para situaciones estresantes a corto plazo, si la situación percibida como amenazadora se prolonga en el tiempo, los mecanismos de respuesta, que en un principio eran agudos y podían revertirse, se vuelven crónicos fatigando el organismo y sin posibilidad de descanso para recuperar su funcionamiento normal, lo que provocaría un deterioro paulatino de la salud tanto mental como física (OIT, 2012).

Si trasladamos el proceso al entorno organizacional, el estrés laboral sería consecuencia de la exposición continuada a factores psicosociales adversos; metafóricamente hablando, un mal jefe sería el antes mencionado león de la sabana. De este modo, la percepción de altas exigencias y un escaso control para hacerles frente, mantenido en el tiempo, estaría imposibilitando una necesaria y adecuada recuperación, convirtiéndose en una situación precursora de enfermedades tanto físicas como psíquicas, además de alteraciones del comportamiento y problemas organizativos.

Intentando proporcionar una idea más aproximada acerca del papel que juega el estrés sobre la salud, Martínez-Plaza (2007) describe de forma sencilla y desde una perspectiva médica los mecanismos de respuesta que se activan con el estrés laboral, implicando al sistema

nervioso central, que coordina a su vez los tres sistemas responsables de la homeostasis o equilibrio del organismo (sistema nervioso autónomo, endocrino e inmune). El autor describe las dos formas en las que el estrés puede producir enfermedad:

1) A través de estados afectivos negativos (ansiedad, depresión), consecuencia del estrés mismo, que actúan mediante procesos biológicos hormonales o conductas de riesgo. 2) A través de respuestas de coping: hacer frente, de forma activa, al agente estresor implica un esfuerzo de adaptación. Esto puede conllevar, en el primer caso, conductas de riesgo (como fumar, ingerir alcohol, abuso de fármacos, entre otros) y, en el segundo, respuestas biológicas hormonales. De esta forma podría incidir, de forma unilateral, en el desarrollo de enfermedad (p.3).

Aunque, tal y como se ha mencionado en párrafos anteriores, no existe un consenso sobre la correcta definición de estrés (P. Martín et al., 2003), sobre lo que sí existe común acuerdo es en considerar que la respuesta al estrés supone una modificación sistémica del organismo (McEwen, 2002; Martínez Plaza, 2007, 2008). Concretamente, el estrés laboral sostenido en el tiempo se relaciona con desórdenes musculoesqueléticos, enfermedades cardiovasculares, depresión, alteración de las funciones inmunológicas, absentismo o suicidio, entre algunos de sus muchos efectos (OIT, 2012).

La literatura existente acerca de las consecuencias del estrés laboral en la salud es tan abundante que, para evitar una extensión inapropiada, se ha incluido en la tabla 4 un compendio de los trastornos mayormente asociados, que en ningún caso limitan otros posibles efectos. En líneas generales las investigaciones se han ido encuadrando dentro de trastornos físicos, psicológicos, sociales y organizacionales.

Tabla 4: Trastornos asociados con el estrés
Fuente: Elaboración propia

TRASTORNOS EMOCIONALES	CITAS
Ansiedad, Obsesiones, fobias, fatiga.	P. Anderson, et al., 2000; Cox, et al., 2005; Cox, Griffiths y Rial-González, 2000; De Jonge, et al., 2000; D'Souza, Strazdins, Lim, Broom y Rodgers, 2003; Labrador y Crespo, 1993; Melchior et al., 2007; Nater et al., 2007; Stansfeld y Candy 2006.
Alienación, apatía, descenso autoestima, cambios de humor, desórdenes mentales, irritabilidad, impaciencia, negativismo.	Bonde, 2008; Colligan, Smith y Hurrell; 1977; De Lange, Taris, Kompier, Houtman y Bongers, 2005; Head, Martikainen, Kumari, Kuper y Marmot, 2002; Kendler, Kessler, Neale, Heath y Eaves, 1993; LaMontagne, Keegel, Louie y Ostry, 2010; Nater et al., 2007; Peled, Carmil, Siboni-Samocha y Shoham-Vardi, 2008; Stansfeld, Clark, Caldwell, Rodgers y Power, 2008; Stansfeld, Fuhrer, Shipley, y Marmot 1999; Wilhelm, Kovess, Rios-Seidel y Finch, 2004; Wang, 2008; Woo y Postolache, 2008.
Depresión.	Blackmore et al., 2007; Bonde, 2008; Burgard, Brand y House, 2009; Glass y Fujimoto, 1994; Kendler et al., 1993; Nater et al., 2007; Netterstrom et al., 2008; Perry-Jenkins, Goldberg, Pierce y Sayer, 2007; Pikhart et al., 2004; Stansfeld y Candy, 2006; Tsutsumi, Kayaba, Theorel y Siegrist, 2001; Van der Doef y Maes, 1999; Wang y Patten 2001; Wang, Patten, Currie, Sareen y Schmitz, 2012; Wieclaw et al., 2008.
Suicidio.	Amagasa, Nakayama y Takahashi, 2005; Martínez-Plaza, 2010; Ostry et al., 2007; Tsutsumi, Kayaba, Ojima, Ishikawa y Kawakami, 2007; Woo y Postolache, 2008.

Continuación tabla 4

TRASTORNOS COGNITIVOS	CITAS
Dificultades de memoria, concentración y pensamientos.	Broadbent, Cooper, Fitzgerald y Parkes, 1982; S. Cohen, 1980; J.J. Kim y Diamond, 2002.
Distracciones, disminución de funciones cognitivas tales como corrección de datos, análisis de situaciones complicadas, incremento del tiempo empleado en resolución de tareas, dificultad toma de decisiones, pensamientos obsesivos.	Christianson, 1992; S. Cohen, 1980; Easterbrook, 1959; Horowitz, 1976; Idzikoski y Baddeley, 1983; Keinan, 1987; Larsen, 2001; Mandler, 1982; Pennebaker, 1989; Staw, Sandelands y Dutton, 1981.
Desorientación, bloqueos mentales, preocupación excesiva, baja percepción de la salud.	Labrador, 1992; Pikhart et al., 2001.
TRASTORNOS FISIOLÓGICOS	CITAS
<u>Trastornos musculo-esqueléticos:</u> Tensión muscular, calambres, contracturas, rigidez dorsalgia, lumbalgia, cervicalgia, puede agravar la fibromialgia y el síndrome de fatiga crónico.	Bongers, 2005; Bongers, de Winter, Kompier y Hildebrandt, 1993; Devereux y Buckle, 2000; Devereux, Vlachonikolis y Buckle, 2002; Engstrom, Hanse y Kadefors, 1999; Feuerstein, Shaw, Nicholas y Huang, 2004; Hagen, Magnus y Vetlesen, 1998; Hoogendoorn, van Poppel, Bongers, Koes y Bouter, 2000; M.S. Kerr, 1998; Larsman, 2006; Martínez-Plaza 2008, 2009a, 2009b; Muniamuthu y Raju, 2010; Stansfeld, Bosman, Hemingway y Marmot, 1998; Theorell, 2000; Torp, Riise y Moen, 2001.
<u>Trastornos digestivos:</u> Caries, bruxismo, estomatitis aftosa, globo histérico, pirosis, dispepsia, úlcera péptica, síndrome de colon irritable, enfermedad de Crohn, colitis ulcerosa, aerofagia, náuseas, vómitos.	Abekura, et al., 2011; Ahlberg et al., 2002, 2013; Cameron y Perdue, 2005; Carvalho, Cury y García, 2008; Collins, Barbara y Vallance, 1999; Drossman, 1998; Duffy et al. 1991; Fukudo y Sukuzi, 1987; Goulet, Lund, Montplaisir y Lavigne, 1993; Levenstein, 1998; Levenstein et al., 1996, 2000; Mawdsley y Rampton, 2006; Mayer, 2000a, 2000b; Martínez-Plaza, 2008; Paar, Bezenberger y Lorenz-Meyer, 1988; Pingitore, Chrobak y Petrie, 1991; Rähä, Kempainen, Kaprio, Koskenvuo y Sourander, 1998; Rao, Bhat y David, 2011; Reber, 2012; Sapolsky, 2004; Winocur, Uziel, Lisha, Goldsmith y Eli, 2011.
<u>Patologías cardiovasculares:</u> Dolor precordial, desarrollo de aterosclerosis, taquicardias, hipertensión, angina de pecho, desarrollo de síntomas de la enfermedad de Raynaud y de acrocianosis, palpitaciones, dolor de pecho...	Belkic, Landsbergis, Schnall y Baker, 2004; Bishop et al., 2003; Bunker et al., 2003; Chandola et al. 2008; Collins, Karasek y Costas, 2005; Eller et al., 2009; Everson-Rose y Lewis, 2005; Head et al., 2002; Kamarck, Shiffman, Sutton-Tyrrell, Muldoon y Tepper, 2012; Karasek, 1979; Karasek y Theorell, 1990; Kivimäki et al., 2002; Kivimäki et al., 2006; Kristensen, 1996; Kuper y Marmot, 2003; Landsbergis et al., 2001, 2003; Lucini, Riva, Pizzinelli y Pagani, 2007; Marmot, Bosma, Hemingway, Brunner y Stansfeld, 1997; Martínez-Plaza 2008; Nater et al., 2007; Schnall, Belkinc, Landsbergis y Baker, 2000; Schnall, Landsbergis y Baker, 1994; Shima y Satoh, 2006; Siegrist, 1996b; Stansfeld y Marmot, 2002; Steenland et al., 2000; Vrijkotte, van Doornen y de Geus, 2000; C. Williams, 1998; R.B. Williams, 2003.
Infarto	Appels, 1993; Appels, Kop y Schouten, 2000; European Heart Network, 1998; Kivimäki et al., 2002; Kristensen, 1996; Theorell y Karasek, 1996.

Continuación tabla 4

<u>Sistema endocrinológico:</u>	Daleva, Piperova-Dalbokova, Hadjiolova y Minchevaet, 1982; Djindjic, Jovanovic, Djindjic, Jovanovic y Jovanovic, 2012; Fröberg, Karlsson, Levi, Lidberg y Seemanet, 1970; Head et al., 2002; Lipton 1976; Martínez-Plaza, 2008.
Hipertiroidismo, aumento triglicéridos y colesterol, aumento glucosa en sangre, diabetes tipo 1 y 2, síndrome metabólico dislipidemias, hipertensión, diabetes mellitus, aumento secreción hormonal.	
<u>Sistema respiratorio:</u>	S. Cohen, Hamrick y Rodríguez, 2002; Martínez-Plaza, 2008; O'Leary, 1990; Peters et al., 1999; Smyth, Soefer, Hurewitz, Kliment y Stone, 1999.
Sensación de ahogo, hiperventilación, disnea, asma bronquial, jadeo laríngeo, rinitis alérgica, agrava síntomas EPOC.	
<u>Sistema dermatológico:</u>	Kiecolt-Glaser, Marucha, Malarkey, Mercado y Glaser, 1996; Martínez-Plaza, 2001, 2008; Park 1998.
Psoriasis, prurito, dermatitis atópica, alopecia, neurodermitis, eczema, hipersudoración, cicatrización lenta de heridas	
<u>Sistema inmunitario:</u>	Cooper, 1986; Eysenck, 1988; Kiecolt-Glaser, Cacioppo, Malarkey y Glaser, 1992; Martínez-Plaza, 2008; O'Leary, 1990; Peters et al., 1999; Segerstrom y Miller, 2004; Smyth et al., 1999.
Artritis reumatoide, lupus eritematoso, diabetes tipo 1, tiroiditis autoinmune, herpes, aumenta el riesgo de atopía (rinitis alérgica, asma bronquial alérgica, dermatitis atópica, anafilaxia, cáncer	
<u>Afecciones neurológicas:</u>	Martínez-Plaza, 2007, 2008; Nater et al., 2007.
Cefaleas, migrañas, ataques epilépticos, amnesia global transitoria, vértigos.	
<u>Trastornos urinarios:</u>	Martínez-Plaza, 2008.
Prostatosis, incontinencia transitoria, enuresis nocturna.	
<u>Trastornos aparato reproductor:</u>	Martínez-Plaza, 2008; Sheiner et al., 2002.
Amenorrea, síndrome premenstrual, disfunción sexual, infertilidad, impotencia, alteraciones de la libido.	
TRASTORNOS DEL COMPORTAMIENTO	CITAS
<u>Trastornos de la conducta alimentaria:</u>	Anton y Miler, 2005; Behar y Valdés, 2009; Brunner, Chandola y Marmot, 2007; Epell et al., 2000; Fowler-Brown et al., 2009; Freeman y Gil, 2004; Gibson, 2006; Gluck, 2006; Greeno y Wing, 1994; Head et al., 2002; Macht, 2008; Ng y Jeffery, 2003; Sassaroli y Ruggerio, 2005; Steptoe, Lipsey y Wardle, 1998; Torres y Nowson, 2007; Wallis y Hetherington, 2009; Ward y Mann, 2000.
Anorexia nerviosa, bulimia nerviosa, falta de apetito, comer compulsivamente, aumento del consumo de determinados alimentos (comida rápida, grasas saturadas, chocolate).	
Sedentarismo o disminución de ejercicio físico.	Aldana, Sutton, Jacobson y Quirk, 1996; Anton y Miller, 2005; Brisson, Larocque, Moisin, Vexina y Dagenis, 2000; Carbone y Cigrang, 2001; Farmer et al., 1988; Hellerstedt y Jeffery 1997; Johansson, Johnson y Hall, 1991; Ng y Jeffery, 2003; Payne, Jones y Harris, 2002.
<u>Trastornos de sueño:</u>	Akerstedt, 2003, 2006; G. Costa, Ghirlanda, Minors y Waterhouse, 1993; De Lange, Taris, Kompier, Houtman y Bongers, 2003; Del Río, 2006; Ganster, Fox y Dwyer, 2001; Ganster y Rosen, 2013; Harrma, 2006; E. J. Kim y Dimsdale, 2007; Monk y Folkard, 1985; Nordin, Knutsson, Sundbom y Stegmayr, 2005; Prospéro, Méndez, Ruiz, Alvarado y León, 2011; Sapolsky, 2004; Sonnentag, Perrewé y Ganster, 2009.
Insomnio, alteración ritmos circadianos.	

Continuación tabla 4

Tartamudeo, temblores, hablar rápido, risa nerviosa, falta de aire, mareos.

Cano-Vindel y Miguel-Tobal, 1999.

Abuso de sustancias:

Fármacos, alcohol, tabaco y drogas, recaídas en periodos de abstinencia.

al' Absi, Carr y Bongard, 2007; Ames, Cunradi, Moore y Duke, 2009; Ames y Delaney, 1992; Ames, Grube y Moore, 2000; Anton y Miller, 2005; Bacharach, Bamberger y Sonnenstuhl, 2002; Bennett y Lehman 1998; Bennett, Paterson, Reynolds, Wiitala y Lehman, 2004; Bilkei-Gorzo et al., 2007; Bray, Fairbank y Marsden, 1999; Brisson et al., 2000; Carreño, Medina-Mora, Martínez, Juárez y Vázquez, 2006; Charro, 2012; Colell, Sánchez-Niubò, Domingo-Salvany, Delclós y Benavides, 2014; Colell, Sánchez-Niubò, Benavides, Delclós y Domingo-Salvany, 2014; Cunradi, Todd, Duke y Ames, 2009; De Lange et al., 2003; Gianoulakis, 1998; Green y Johnson 1990; Head et al., 2002; Hellerstedt y Jeffery 1997; Johansson et al., 1991; Kishchuk et al. 1994; Kouvonen, Kivimaki, Virtanen, Pentti y Vahtera, 2005; Marín, 2013; Ng y Jeffery, 2003; Ochoa y Madoz, 2008; Rodríguez et al, 2008; Stansfeld et al., 1999; Towers, Kishchuck, Sylvestre, Peters y Bourgault, 1994; Trice y Roman, 1978.

ORGANIZACIÓN**CITAS**

Impuntualidad, absentismo, escaso rendimiento.

Bardera et al., 2002; Elovainio, Kivimaki, Vahtera, Virtanen y Keltikangas-Jarvinen, 2003; González-Romá et al., 2005; Hart y Wickens, 1990; J.H. Kerr, 1996; Levi, Frandenhaeuser y Gardell, 1986; Luceño 2005; Luceño y Martín, 2005; Manning y Osland 1989; Marmot, Feeney, Shipley, North y Syme, 1995; J. Martín et al., 2007; J.M. Molina, 2010; Moncada 1998; Niedhammer, Goldberg, Leclerc, Bugel y David, 1998; Óscar, González-Camino, Bardera y Peiró, 2003; Sutherland y Cooper, 1988; Tomás, Rodrigo y Oliver, 2005; Undén 1996.

Accidentabilidad, lesiones.

Bardera et al., 2002; Cano-Vindel et al., 2011; Carreño, et al., 2006; Glasscock, Rasmussen, Carstensen y Hansen, 2006; Lauridsen y Tonnesen, 1990; Murray, Fitzpatrick y O'Connell, 1997; Hart y Wickens, 1990; Rundmo, 1994; Stager et al., 1989.

Rotación laboral e intentos de abandono laboral.

Kohan y Mazmanian, 2003.

Nota: Las consecuencias del estrés aquí recogidas no son excluyentes entre sí.

Todas las evidencias científicas recopiladas en la tabla 4 forman parte de una lista que lejos de detenerse va *in crescendo*, lo que sirve de referencia para hacernos una idea de la magnitud de la gravedad asociada a este riesgo. “Necesitamos entender que el estrés en el trabajo es un fenómeno que no tiene olor ni color, pero sus efectos negativos en la salud y el bienestar tanto de las personas como de las organizaciones es devastador” (Dolan, García y Diez Piñol, 2005, p. xv). Precisamente, debido a esta nefasta invisibilidad, se está empezado a acuñar el término de *asesino silencioso* como analogía de estrés (Usha, 2016).

1.4.2. Síndrome de *Burnout*

El síndrome de desgaste profesional, también conocido como síndrome de quemarse en el trabajo (SQT) o *Burnout* es un tipo específico de estrés laboral. Gil-Monte (2003) en una revisión encontró hasta 19 denominaciones distintas en español para hacer referencia a este riesgo; no obstante, la denominación de *burnout* es la más extendida entre los investigadores.

Esta expresión fue utilizada por primera vez en 1974 por el psiquiatra Herbert Freudenberger quien la definió como “sensación de fracaso y una existencia agotada o gastada que resultaba de una sobrecarga por exigencias de energías, recursos personales o fuerza espiritual del trabajador” (p. 160).

Freudenberger, desde un enfoque clínico, describió los diversos signos físicos y conductuales que padecían los voluntarios que trabajaban con toxicómanos en la *Free Clinic* de Nueva York. Observó que transcurridos entre uno y tres años la mayoría de voluntarios mostraban una paulatina pérdida de energía que derivaba en agotamiento y apatía, manifestando alteraciones de conducta, tales como frustración, paranoia, depresión, falta de sensibilidad, irritabilidad y cinismo hacía las personas que prestaban sus servicios (Freudenberger y Richelson, 1980).

Posteriormente, Maslach y Jackson (1981a), fruto del resultado empírico de sus investigaciones, entienden el *burnout* como “un síndrome tridimensional caracterizado por agotamiento emocional, despersonalización y reducida realización personal” (p. 3). Esta es, a día de hoy, la definición más considerada por la comunidad científica (Pocinho y Perestrelo, 2011). A su vez, Maslach es la autora del instrumento principal para la medición y estudio del síndrome: el *Maslach Burnout Inventory* (MBI, Maslach y Jackson, 1981b).

Existen muchas otras aproximaciones teóricas al respecto. Aún sin llegar a un consenso unánime, la mayoría coincide al identificar el SQT como la respuesta prolongada a estresores crónicos derivados de las relaciones interpersonales, con consecuencias negativas para el individuo y para la organización en la que trabajan (Maslach, 2009; Mingote, Moreno y Gálvez, 2004). Se distinguen dos líneas de estudio (Gil-Monte y Peiró, 1997): (a) una perspectiva clínica, donde el *burnout* es entendido como un estado personal, un conjunto de síntomas, al que llega el individuo como resultado del estrés laboral; y (b) un enfoque psicosocial –el que cuenta con más adeptos– entendido como un proceso evolutivo, producto de la interacción entre las características personales del trabajador y su entorno laboral, identificado por los tres dimensiones descritas por Maslach y Jackson (1981a). En base a esos planteamientos, Gil-Monte (2003) realiza una definición muy ilustrativa del síndrome:

El síndrome de quemarse por el trabajo se define como una respuesta al estrés laboral crónico caracterizada por la tendencia de los profesionales a evaluar negativamente su habilidad para realizar el trabajo y para relacionarse con las personas a las que atienden, por la sensación de estar agotados a nivel emocional, por el desarrollo de sentimientos negativos, y de actitudes y conductas de cinismo hacia las personas destinatarias del trabajo, que son vistas de manera deshumanizada debido al endurecimiento afectivo del profesional. El síndrome de quemarse por el trabajo no debe identificarse con estrés psicológico, sino que debe ser entendido como una respuesta a fuentes de estrés crónico (estresores) que se originan en la relación profesional-cliente, y en la relación profesional-organización (p. 183).

En un principio, el síndrome se ha relacionado con profesiones que implican trato directo con personas (Ganster y Schanbroeck, 1991; García-Izquierdo, 1991a, 1991b; Moreno, Oliver y Aragoneses, 1991), presentando mayor prevalencia en los profesionales del sector servicios dedicados a la ayuda continua a los demás, especialmente en las ramas de sanidad y educación (Gil-Monte, 2009; Grau, Flichtentrei, Suñer, Prats y Braga, 2009; Mingote et al., 2004). Sin

embargo, Pines y Aronson (1988) no limitan el SQT a las profesionales de servicios y definen el *burnout* como “el estado de agotamiento mental, físico y emocional, producido por la involucración crónica en el trabajo en situaciones con demandas emocionales” (p. 9). Siguiendo esta directriz, en 1996, los autores de la tercera edición del MBI (Maslach, Jackson y Leiter, 1996), reemplazaron la dimensión de despersonalización por la de cinismo, consiguiendo hacer extensible el estudio del *burnout* a todo tipo de profesiones como deportistas (Carlin y Garcés de los Fayos, 2010) o estudiantes (Caballero, Hederich y Palacio, 2010; I.M. Martínez y Salanova, 2003; Rosales y Rosales, 2013).

Las repercusiones del SQT pueden llegar a ser graves al afectar a la esfera personal, laboral y familiar. A nivel individual, cuando los mecanismos de afrontamiento y autoprotección fallan, se puede producir un deterioro que afecta a los planos cognitivo, emocional, actitudinal, psicosomático y social (Carlin y Garcés de los Fayos, 2010; Fidalgo, 2006a; Gil-Monte y Peiró, 1997; Rosales y Rosales, 2013). La sintomatología asociada al síndrome abarca un gran número de síntomas, más de 100 según la revisión de Fidalgo (2006b) en la NTP 705. En la siguiente tabla se recogen algunos de los síntomas más destacados.

Tabla 5: Trastornos asociados al *Burnout*

Fuente: Caballero, et al. (2010); Carlin y Garcés (2010), Fidalgo (2006b), Gil-Monte y Peiró (1997) y Quiceno y Vinaccia, (2007).

TRASTORNOS	
EMOCIONALES	Agotamiento emocional, ansiedad, apatía, depresión, desesperanza, desilusión, desorientación, distanciamiento emocional, expresiones de hostilidad, impaciencia, indefensión, intolerancia, irritación, pesimismo, sentimientos de alienación, sentimientos de impotencia, sentimientos de omnipotencia, sentimientos de soledad, supresión de sentimientos.
COGNITIVOS	Baja realización personal laboral, criticismo generalizado, desaparición de expectativas, desorientación cognitiva, distracción, fracaso profesional, modificación del auto-concepto, pérdida de autoestima, pérdida de la creatividad, pérdida de significado, pérdida de valores.
DE ACTITUD Y CONDUCTA	Abuso de alcohol y cafeína, abuso de barbitúricos, abuso de tabaco y drogas, actitudes negativas (verbalizar), apatía, auto sabotaje, cinismo, comportamiento suspicaz y paranoide, conductas inadaptadas, conductas obsesivo compulsivas, desorden bipolar, desorganización, desvalorización, evitación de decisiones, evitación de responsabilidades, hostilidad, inflexibilidad y rigidez, incapacidad de concentración, mala praxis.
PSICOSOMÁTICOS	Alteraciones del sistema nervioso, alteraciones digestivas (úlceras, gastritis, diarrea, etc.) alteraciones respiratorias, altos índices analíticos (colesterol, triglicéridos, glucosa, ácido úrico, etc.), cambios de peso, cansancio, disfunciones sexuales, dolores de cabeza, espalda y cuello, dolores musculares, enfermedades cardiovasculares, insomnio, pérdida de apetito, quejas psicosomáticas.
SOCIALES	Agresividad hacia clientes, aislamiento y sentimientos de fracaso, conflictos interpersonales, evitación de contactos, formación de grupos críticos, irritabilidad, mal humor familiar, problemas de pareja, quejas constantes, superficialidad.
PARA LA ORGANIZACIÓN	Absentismo laboral, aumento conflictos, baja satisfacción laboral, desinterés por realizar las tareas, deterioro de la comunicación, disminución capacidad de trabajo, incremento de quejas de usuarios o clientes, indiferencia hacia el trabajo, intenciones de abandono, mayor desmotivación, minoración de la calidad de vida laboral, relaciones interpersonales frías y distantes

Toda esta sintomatología conforma un cuadro clínico que, de no corregirse, iría evolucionado con el paso del tiempo. Aunque determinadas características personales (ej., edad, género o estilos de afrontamiento) puedan actuar como factor protector, las consecuencias tanto a nivel individual como de la organización son tan importantes que, para poder paliar sus efectos -a menudo devastadores-, es necesario establecer medidas de prevención que traten de identificar sus causas y poder así eliminarlas o reducirlas.

1.4.3. Mobbing

El acoso laboral, fundamentalmente conocido en España por el anglicismo *mobbing*, es una forma más de estrés laboral, con la peculiaridad de que su origen no se encuentra en el desempeño del trabajo, sino que es causado por las relaciones interpersonales dentro de la organización (Dollard, Skinner, Turkey y Bailey, 2007; Martín-Daza, Pérez-Bilbao y López; 1998; Mayoral y Espluga, 2010). Por tanto, es la mala calidad de las relaciones laborales la que afecta al bienestar de los trabajadores, transformándose en un factor de riesgo. Moreno (2004, p. 3) lo considera “como una de las formas más severas de estrés psicosocial en el trabajo”. Su importancia es tal que multitud de publicaciones se refieren a él como *la nueva plaga del siglo XXI* (Acosta, 2001; Marcos y Velázquez, 2011; Piñuel, 2001; Piñuel y Oñate, 2006; Verona y Santana, 2012; Vivas, Delgadillo y Mercado, 2008).

En relación al vocablo *mobbing*, éste fue introducido en las ciencias sociales por el etólogo Konrad Lorenz (1963, 1965) para referirse a las conductas animales donde los miembros débiles se aliaban para atacar a un individuo más fuerte de la misma especie. En 1972, Peter Paul Heinemann, médico sueco, utilizó la misma expresión para describir el comportamiento beligerante de un grupo de niños, en oposición a otro niño de la escuela. Sin embargo, se atribuye a Leymann (1986) la aplicación del fenómeno *mobbing* en la esfera de la psicología laboral, al reproducir el término para definir un tipo de violencia de grupo, similar al escolar, aunque más sofisticado, en un contexto laboral y entre adultos. De este modo, Leymann es considerado como el padre de las investigaciones sobre *mobbing* (Piñuel, 2008), y a su vez, es el creador del primer cuestionario para evaluar el *mobbing* en *Leymann Inventory of Psychological Terrorization*, LIPT (Leymann (1992). A día de hoy, su definición es una de las más citadas (Leymann, 1996a):

La interacción social, a través de la cual, una persona (rara vez más), es atacado por uno o más (rara vez más de cuatro) individuos casi diariamente y durante periodos de varios meses, llevando a la persona a una posición casi de indefensión, con un alto riesgo potencial de expulsión (p. 168).

Desde entonces, varios investigadores han realizado aportaciones acerca del acoso sin que se haya logrado reunir un criterio comúnmente aceptado, utilizando aproximaciones distintas al *mobbing* tales como:

- *Bullying (in the workplace)*: comportamientos ofensivos con el fin de intimidar, degradar y humillar a una persona (Di Martino, Hoel y Cooper, 2003; Hoel, Cooper, y Faragher, 2001). En comparación con las técnicas sutiles y sofisticadas del *mobbing*, el *bullying* estaría más relacionado con la violencia física (Leymann, 1996b).

- *Harassment*: es el término escandinavo utilizado para referirse al acoso laboral, se caracteriza por la persistencia de una persona en provocar y atormentar a otra que no es capaz de defenderse (Björkqvist, Österman y Hjelt- Bäck, 1994).
- *Emotional abuse* [abuso emocional]: se utiliza como sinónimo de *mobbing* en Canada y Estados Unidos (Keashly, 1998).
- *Ijime*: es el término japonés utilizado para definir el acoso laboral (Hirigoyen, 2013a, 2013b; López y Vázquez, 2005).
- *Petty tyranny* [pequeña tiranía]: trata sobre la tiranía de los líderes que utilizan el poder para despreciar a sus subordinados (Ashforth, 1994).
- *Psicoterror laboral*: este concepto fue usado por Leymann (1990a) en los inicios de sus investigaciones sobre *mobbing*, utilizándolo como sinónimo.
- *Scapegoating* [chivo expiatorio]: la persona que carga con las culpas de otros (Thylefors, 1987, citado por Sáez y García-Izquierdo, 2009).
- *Victimization* [victimización]: este concepto se refiere al proceso de victimización que sufren los trabajadores objeto de acoso laboral (Einarsen y Raknes, 1997).
- *Whistleblower* [el que se va de la lengua]: se refiere a las personas que sacan a la luz y alertan sobre los comportamientos irregulares de sus compañeros de trabajo, convirtiéndose por ende en objeto de represalias por parte de otros miembros de la organización. (Hirigoyen, 2013a, 2013b; López y Vázquez, 2005).
- *Workplace trauma*: este concepto describe el punto de vista del trabajador, víctima de continuos comportamientos por parte de superiores o compañeros con objeto de desintegrar su autoestima (Wilson, 1991, citado por Sáez y García-Izquierdo, 2009).

Todos estos conceptos hacen referencia, con ciertas matizaciones, a un maltrato deliberado, continuado y constante sobre una determinada persona, utilizando comportamientos negativos de hostilidad (Meseguer, Soler y García-izquierdo, 2007)

Circunscribiéndonos a la esfera laboral, Leymann (1996a) estableció que para poder hablar de *mobbing* tendrían que repetirse, al menos una vez a la semana y durante un período mínimo de seis meses, alguno de los 45 tipos de conducta descritos en *Leymann Inventory of Psychological Terrorization*, LIPT (ej., *el acosador no permite la libre comunicación de la persona acosada; se le interrumpe continuamente cuando habla; se le impide expresarse; es objeto de avasallamiento, gritos o insultos en voz alta.*). Siguiendo a Bartolomé, Fernández, Prieto y Sánchez (2007, p. 348):

Estos comportamientos deliberados van dirigidos a conseguir la anulación profesional y personal de la víctima. Con ellos intentan que la víctima no pueda expresarse, intentan reducir y destruir sus relaciones sociales, desprestigiarlo a todos los niveles y, si es necesario, llegan a atentar contra su salud física y psíquica, de manera que logran, en todos los casos, mermar la calidad de vida de la persona contra la cual atentan.

Por lo tanto, son elementos distintivos de *mobbing* la sistematización de las conductas de acoso, su persistencia temporal y la clara intención de dañar. Llegados a este punto, cabe especificar que existen situaciones que se confunden con el *mobbing* y no pueden ser

catalogadas como tal. Hirigoyen (2013a) realiza una clara diferenciación entre condiciones de acoso laboral y otras situaciones de sufrimiento laboral como por ejemplo la presión en el trabajo, las desavenencias, las malas condiciones de trabajo, las críticas constructivas, las agresiones esporádicas o el conflicto.

El acoso laboral es un atentado contra la dignidad de los trabajadores, que genera sufrimiento a quien lo padece. En palabras de Leymann (1996a, p. 26) “el *Mobbing* es un proceso de destrucción; se compone de una serie de actuaciones hostiles que, tomadas de forma aislada, podrían parecer anodinas, pero cuya repetición constante tiene efectos perniciosos”. Según recogen Trujillo, Valderrabano y Hernández (2007), el ahogo, las palpitaciones, el cansancio, el pánico o la pérdida de memoria son algunos de los síntomas físicos que sienten las víctimas al poco tiempo de comenzadas las conductas de acoso. En la tabla 6 se describen algunas de las consecuencias que pueden sufrir las víctimas de *mobbing* según diversos autores (Leymann, 1996a; Piñuel, 2001, 2002, citados por M. Martínez, Irurtia, Martínez, Torres y Queipo, 2012).

Tabla 6: Trastornos asociados al Mobbing.

Fuente: M. Martínez et al. (2012)

TRASTORNOS	
PSICOLÓGICOS	Sentimientos de fracaso e impotencia, frustración, baja autoestima, apatía, problemas de concentración y atención, ansiedad, depresión, trastornos paranoides, comportamientos suicidas, pérdida de memoria, aislamiento, incapacidad de mostrar alegría, irritabilidad, hipersensibilidad a la crítica, trastorno de estrés postraumático, drogodependencia y otras adicciones
FÍSICOS	Insomnio, pesadillas, ataques de pánico sudoración, palpitaciones, fatiga crónica, dolor muscular, entumecimiento de extremidades, pérdida de apetito
PARA LA ORGANIZACIÓN	Jubilación forzosa, despido, absentismo, rotación del personal, baja voluntaria o enfermedad y pérdidas económicas

En definitiva, las consecuencias del acoso son graves y no sólo para la salud de los afectados, sino que también repercuten negativamente en su ámbito socio-familiar y en la propia organización, pues el *mobbing* puede suponer largos periodos de absentismo laboral y descenso de la productividad (González-Trijueque, et al., 2012; Leymann, 2009; Trujillo et al., 2007).

Las investigaciones sobre acoso laboral de los últimos años desvelan que el porcentaje de afectados por *mobbing* se ha incrementado notablemente, llegando López-Cabarcos y Vazquez-Rodríguez (2005) a referirse a él como el cáncer laboral del nuevo milenio, término nada exagerado si hacemos referencia a los datos del Barómetro CISNEROS XI del año 2009, donde más de 2,5 millones de trabajadores en España son víctimas de *mobbing*, un 43% más de afectados que en el año 2006 (Verona y Santana, 2012). Para paliar estos datos, sería imprescindible que las organizaciones, como responsables de evitar la materialización de los riesgos laborales, implantaran medidas preventivas adecuadas como son las evaluaciones de riesgos psicosociales puesto que la organización del trabajo, el estilo de liderazgo, la cultura organizacional, el clima laboral y el entorno socio-económico son, entre otros, algunos de los factores psicosociales predictores de *mobbing* (González-Trijueque y Delgado, 2008).

1.5. PANORAMA ACTUAL Y PERSPECTIVAS DE FUTURO DE LOS RIESGOS PSICOSOCIALES

Como ha quedado patente, los factores psicosociales están estrechamente relacionados con la organización de las empresas. Ésta lleva décadas experimentando continuos reajustes, propiciados por la globalización, la incorporación de nuevas tecnologías, el predominio del sector servicios o la flexibilización laboral, entre otros (Benach y Muntaner, 2007; Blanch, 2003; Gil-Monte, 2009; Leka y Jain, 2010; Luceño, Martín, Rubio et al., 2008; Peiró, 2000, 2001, 2008; Peiró, Yeves y Lorente, 2013; Porthé et al., 2009; Useche, 2002). Esta transformación ha aportado nuevas peculiaridades al mundo laboral, sobre todo en forma de factores de riesgo como el aumento del trabajo mental en detrimento del físico, los distintos modos de relación entre empleado y empleador, las desigualdades en el salario, las malas relaciones interpersonales, las tareas rutinarias, la presión de tiempo, las nuevas formas de contratación, la falta de promoción, la inseguridad en el empleo y un largo etcétera (Luceño, Martín, Rubio, et al., 2008; P. Martín, Salanova y Peiró, 2003; Peiró, 2001, 2004, 2008, Porthé et al., 2009; Tomasina, 2012).

Este aumento de la exposición a los factores de riesgo psicosocial lo que hace es poner a prueba la capacidad de respuesta de los trabajadores, que ven incrementadas sus exigencias laborales, propiciando situaciones de tensión laboral (EU-OSHA, 2009; Fernández-López, Siegrist, Rödel y Hernández-Mejía, 2003; Garrido-Pinzón, Uribe-Rodríguez y Blanch, 2011; OIT, 1984, 1986; Peiró, 2000, 2001, 2008; C. Sánchez y Conde, 2008; L. Sánchez et al., 2008; Sinclair, Sears, Probst y Zajack, 2010). Esta tensión, además de influir en la aparición de insatisfacción laboral, (1) puede ser origen de multitud de enfermedades (Luceño, Martín, Díaz y Rubio, 2008; Moncada, Llorens, Gimeno y Font, 2007; Theorell, 2012); (2) incrementa el riesgo de accidentes (Moreno y Báez, 2019); y (3) repercute de manera negativa en los resultados económicos de la empresa (Moreno y Báez, 2010).

Si las circunstancias ya eran de por sí preocupantes y desalentadoras antes de poner pie en escena la reciente crisis económica mundial, las consecuencias de ésta han desbaratado todo el panorama. La hecatombe financiera ha provocado que hayan ido en aumento las exigencias cuantitativas y que, a su vez, haya disminuido la percepción de autonomía de los trabajadores al haberse intensificado la carga y la jornada de trabajo. Además de ello, se desempeñan funciones de inferior nivel al de la categoría profesional reconocida, ocasionando un descenso de la autoestima. Se añaden las bajas remuneraciones, la paulatina desaparición del empleo estable en favor del empleo temporal y el considerable aumento de las tasas de desempleo, que llevan aparejados una mayor exposición al factor de inseguridad y un descenso significativo del apoyo social (Luceño, Martín, Rubio, et al., 2008; Senet, 2006; Trillo, 2009). Al respecto, la Declaración de Edimburgo sobre la promoción de la salud mental y el bienestar en el trabajo (European Network for Workplace Health Promotion, ENWHP, 2011, p. 1) destaca que, en Europa el absentismo, el desempleo y las solicitudes de incapacidad permanente relacionados con el estrés y los problemas de salud mental en el trabajo han ido en aumento, pudiéndose atribuir alrededor del 10% de los problemas crónicos de salud a trastornos mentales y emocionales.

Conocedor de esta realidad, el Observatorio Europeo de Riesgos ha publicado acerca de los factores psicosociales que éstos serán unos de los riesgos laborales más importantes en los próximos años para los trabajadores europeos y que podrán conducir a problemas de salud como

depresión, estrés, ansiedad o problemas de sueño, entre otros. Asimismo, agrupa en cinco bloques los factores que contribuirán a acentuar dichas enfermedades, y que serán las nuevas formas de contratación laboral e inseguridad en el puesto de trabajo, las fuertes exigencias emocionales en el trabajo, el desequilibrio entre la vida profesional y personal, la intensificación del trabajo y el envejecimiento de la población en edad de trabajar (EU-OSHA, 2009).

Estos argumentos no iban desencaminados, pues el estrés es causa de entre un 50% y un 60% de las bajas laborales (Takala, 2010). Según EU-OSHA (2014a), que se hace eco de los datos obtenidos en el *Sondeo de opinión panauropeo sobre seguridad y salud en el trabajo* (EU-OSHA, 2013), un 51% de los trabajadores considera que el estrés es frecuente en sus lugares de trabajo, y 4 de cada 10 trabajadores creen que el estrés no se aborda de forma correcta en las empresas en las que trabajan. Las causas más frecuentemente mencionadas fueron la organización del trabajo o la precariedad laboral (72%), la larga jornada laboral o una excesiva carga de trabajo (66%) y el sentirse intimidado o acosado en el trabajo (59%). La lenta recuperación económica de países como España hace que la precariedad y la pobreza laboral sigan, lamentablemente, al orden del día (Eurostat, 2017).

Este panorama, nada halagüeño, se confirma por la comparativa de los resultados de la Encuesta Nacional de Condiciones de Trabajo (ENCT) realizada antes y durante el periodo de crisis financiera. Así, al confrontar los resultados de las dos últimas encuestas, relativas a los años 2007 (ENCT VI) y 2011 (ENCT VII), se refleja un claro empeoramiento de los datos referentes al ámbito psicosocial. En la tabla 7 se presenta a modo de ejemplo algunos porcentajes de exposición de ambas encuestas, permitiendo observar el agravamiento de los resultados.

Tabla 7: Comparación de porcentajes de exposición a factores psicosociales entre ENCT VI y ENCT VII
Fuente: Elaboración propia

Percepción de:	ENCT VI (2007)	ENCT VII (2011)
Incremento del nivel de exigencia	67.70	77.60
Tener mucho trabajo y sentirse agobiados	20.30	24.10
Tener que trabajar muy rápido	44.00	46.00
Atender varias tareas al mismo tiempo	41.20	45.30
Causas de accidentes:		
Trabajar muy rápido	9.70	18.20
Cansancio o fatiga	3.50	13.60
Exceso de horas continuadas de trabajo	0.90	3.10

Otro dato significativo que se desprende de ambas encuestas es la información recogida sobre las molestias musculoesqueléticas sufridas por los trabajadores, que son trastornos muy relacionados con los factores de riesgo psicosocial. Se observa que el porcentaje de trabajadores que indicaron sufrir dicha dolencia en el 2011 fue de 77.50%, lo que constituye 3.8 puntos por encima de la anterior encuesta (ENCT VI, 2007), que alcanzó el 73.70%.

Como consecuencia, lejos de enmendar los riesgos ya existentes se está propiciando su agravamiento (ej., estrés laboral) y la aparición de nuevos riesgos como el *burnout*, el acoso psicológico en el trabajo, o la violencia (EU-OSHA, 2009; Luceño, Martín, Rubio et al., 2008;

Moncada, et al., 2011; Peiró, 2004; WHO, 2010), lo que nos lleva a colegir que los riesgos psicosociales son un problema creciente para la Salud Pública.

Corroborando esta afirmación se encuentran los informes y encuestas realizados a lo largo de los años por diversos organismos. Entre ellos Regus (2009). Los resultados de su encuesta, a 16000 trabajadores de todo el mundo, manifiestan que el 58% de compañías experimentó un incremento significativo de estrés laboral en los últimos dos años. En España, la última Encuesta sobre Calidad de Vida en el Trabajo (ECVT, 2010) publicada por el Ministerio de Trabajo e Inmigración, arroja datos reveladores, puesto que un 49.90% de trabajadores manifiestan niveles altos o muy altos de estrés, en oposición al 17.40% que presentan niveles bajos o muy bajos. De igual modo, en la V Encuesta Europea sobre Condiciones de Trabajo (5EWCS, Eurofound, 2012), un 20% de los trabajadores encuestados declararon que sus condiciones de trabajo suponían un riesgo para su salud mental. Otro organismo que se ha pronunciado al respecto es la OMS, quien realiza un fatídico pronóstico para el año el 2020, augurando que la depresión, principal síntoma del estrés, será la segunda causa más importante de incapacidad en el mundo: A día de hoy más de 150 millones de personas en el mundo padecen depresión por motivos laborales (OIT, 2012). La primera causa de muerte relacionada con el trabajo le corresponde a la enfermedad cardiovascular (ECV), estrechamente asociada al estrés (B.E. Cohen, Edmondson y Kronsich, 2015), con una estimación de 17,5 millones de muertes en el año 2012, cifra que representaba más del 30% del total de fallecimientos (OIT, 2016).

No hay que olvidar que los riesgos psicosociales, aparte de causar los consabidos daños a la salud de los trabajadores, también ocasionan graves perjuicios en la economía de las organizaciones. Sin ir más lejos, se estima que los costes de los trastornos de salud mental en Europa suponen 240 millones de euros anuales, y que la disminución de la productividad asociada al absentismo laboral origina gastos de 136 millones de euros (ENWHP, 2011). Otras fuentes más recientes (MATRIX, 2013) apuntan cifras más elevadas, achacando 617 000 millones de euros anuales a los costes de la depresión de origen laboral y de 272 000 millones de euros al absentismo y presentismo. En lo concerniente a España, el estudio elaborado por el Observatorio de Riesgos Psicosociales de UGT (ORP-UGT, 2013), cifra el coste ocasionado por trastornos y enfermedades mentales atribuibles a las condiciones laborales en 170 960 millones de euros, y entre 150 y 372 millones de euros el gasto sanitario destinado a esos trastornos y comportamientos mentales relacionados con la esfera laboral. Es difícil hallar estimaciones que se aproximen, no obstante, sean unas u otras las más certeras, la realidad es que estamos ante un problema de gran magnitud.

Datos tan funestos hacen inevitable preguntarse si se están evaluando correctamente los riesgos psicosociales y si las medidas que se utilizan son las adecuadas. Resulta obvio que algo no se está haciendo bien. Lo cierto es que, a pesar de su obligatoriedad y de las evidencias sobre sus efectos negativos, las iniciativas preventivas han dirigido mayor esfuerzo y recursos a investigar los aspectos físicos en detrimento de los psicosociales (Cifre, et al., 2011; ESENER, 2010; ESENER-2 2015; M. García y Olarte, 2010; Merino, 2013). Como muestra se hallan los resultados de la Encuesta Nacional de Gestión Preventiva en las Empresas (INSHT, 2009a), en la que algunos de los aspectos más evaluados por las organizaciones han sido la seguridad de los equipos de trabajo, los esfuerzos físicos y el diseño del puesto trabajo, quedando los factores

psicosociales en un segundo plano. Para reforzar estos postulados se pueden consultar los resultados de la ENCT VII (2011), donde se comprueba que los estudios referidos a factores psicosociales (ej., pausas, horarios, estrés) ocupan tan sólo un 28.80%, frente al 51.70% dedicado a posturas, esfuerzos físicos y movimientos repetitivos; el 41.90% referente al ruido; el 40.60 % de estudios sobre la seguridad en las instalaciones o el 35.50% que ocupan la seguridad de máquinas, equipos y material del puesto de trabajo. Datos más recientes, extraídos de la encuesta ESENER-2 (2015), permiten vislumbrar un escenario donde los riesgos psicosociales siguen siendo meros figurantes a la hora de incluirlos dentro de las evaluaciones de riesgos (tabla 8).

Tabla 8: Ranking de los factores de riesgos incluidos en las evaluaciones de riesgos laborales
Fuente: Adaptación a partir de ESENER-2, 2015

Factores de riesgo evaluados	España %	EU %
Sustancias químicas o biológicas	92.30	89.60
Posturas de trabajo, carga física, movimientos repetitivos	91.60	74.70
Seguridad máquinas, equipos e instalaciones	87.10	83.80
Exposición ruido, vibraciones, calor o frío	62.10	70.90
Aspectos organizativos (ej., horarios, pausas, turnos)	59.90	65.40
Relación entre trabajador y supervisor	54.40	53.50

Nota: Se ha destacado en gris lo relativo a factores de riesgo psicosocial

Para refrendar estos porcentajes, resulta revelador que los expertos menos solicitados en España por los servicios de prevención, tanto internos como externos, sean los referentes al área psicosocial (15.70%), en contraposición con los médicos del trabajo (79%) y los expertos en prevención de accidentes laborales (76%) que son los más requeridos (ESENER-2, 2015). En definitiva, son más de 20 años los que han pasado desde la entrada en vigor de la LPRL y a la vista de la escasez de aspectos organizativos evaluados, tal y como muestran las estadísticas, queda patente su reiterado incumplimiento.

Que los riesgos psicosociales sean los menos valorados es una realidad por todos conocida, incluso avalada por el propio Comité de Altos Responsables de la Inspección de Trabajo (SLIC), quienes sabedores de ello pusieron en marcha una campaña europea de riesgos psicosociales (SLIC, 2012). En España, la ITSS fue la encargada de llevarla a cabo a través de la instrucción 2/2012 SLIC, y los resultados revelaron que el 38.50% de empresas participantes carecía de evaluaciones de riesgos psicosociales, y de ellas el 88% necesitaron requerimientos de mejora al ser inadecuadas. Si tenemos en cuenta que es muy común la falta de convenios entre empresas y servicios de prevención para la gestión de estos riesgos, y que a menudo las evaluaciones iniciales de riesgos laborales dejan fuera la parcela referente a *lo psicológico*, no puede sorprender que 2/3 de las evaluaciones de riesgos psicosociales hayan sido promovidas a consecuencia de requerimientos de la inspección laboral (Molina, 2010).

La escasez de estudios sobre riesgos psicosociales hace difícil encontrar variedad de fuentes que contribuyan a sostener estos datos. No obstante, lo argumentado hasta ahora es

suficiente para comprender que el ámbito psicosocial es el de menor actividad preventiva, lo que convierte la evaluación y la prevención de estos riesgos en “la asignatura pendiente”, la cual sigue sin ocupar un lugar destacado ni en las agendas políticas ni en la mayoría de las organizaciones (Barreiro, 2006; Cifre et al, 2011; Fernández, R., 2010; M. García y Olarte, 2010; S. Jiménez, 2014; Leka y Jain, 2010; Merino, 2013; C. Molina, 2010; Cifre, Salanova y Rodríguez, 2009).

Proposición 1:

La mayoría de empresas concede más importancia a la evaluación de los riesgos físicos que a la evaluación de los riesgos psicosociales.

Intentando buscar una posible explicación al rol secundario que juegan los riesgos psicosociales en las iniciativas preventivas, se baraja la ausencia de criterios establecidos legalmente. Aunque en nuestro país existe un amplio abanico de imperativos legales, instrucciones y recomendaciones para la evaluación de los riesgos físicos (ej., de seguridad, de higiene o ergonómicos), no ocurre lo mismo con los riesgos psicosociales, ya que España sigue siendo uno de los pocos países de la EU sin normativa específica que aborde estos riesgos (M. García y Olarte, 2010; Cifre, et al., 2009). La falta de un marco legal definido que arroje luz sobre cómo evaluar los riesgos psicosociales encamina a los profesionales a utilizar como orientación los artículos 5.2 y 5.3.d) del RSP, desde donde se insta a utilizar, por este orden: normativa específica, normas UNE, guías del INSHT o de otros organismos competentes y normas internacionales. En ausencia de éstas, como es el caso de dichas evaluaciones psicosociales, se deberán utilizar guías de otras entidades de reconocido prestigio en la materia u otros métodos o criterios profesionales descritos documentalmente, siempre que proporcionen confianza en el resultado. De ese modo, existen a disposición de los profesionales infinidad de herramientas de evaluación de riesgos que afirman ser fiables y válidas, la mayor parte basadas en medidas de autoinforme.

Si a esta falta de legislación y criterios establecidos se suma la notoria invisibilidad de este tipo de riesgos, no es de extrañar que los empresarios aleguen la existencia del doble de carencias para gestionar los riesgos psicosociales (23.20%) que para tratar los riesgos físicos (11%). Las dificultades que entraña la evaluación psicosocial puede ser responsable de que tan sólo un 15% de empresas de más de 10 trabajadores toman medidas para prevenir esos riesgos (ESENER, 2010). A colación, Cifre et al. (2009) recopilan datos de distintas encuestas a empresas sobre prácticas preventivas en distintas regiones de España y los resultados hablan por sí solos: Navarra, con un 10% de centros de trabajo que realizan intervenciones sobre el estrés, es la comunidad con el porcentaje más elevado de intervención en materia psicosocial. En este sentido, un interesante estudio Delphi basado en las opiniones de médicos del trabajo (Igartúa, Juanola, Mariné y Molinero, 2006, citado por I. Cortés et al, 2010), evidenció los principales obstáculos que limitan la gestión de los riesgos psicosociales:

- a) la reticencia por parte de los directivos para actuar sobre los riesgos psicosociales,
- b) la falta de protocolos de uso sencillo, y
- c) la dificultad para relacionar el trastorno mental con el origen laboral.

Con el paso de los años, y a pesar de las campañas de concienciación llevadas a cabo por diversidad de organismos, los datos siguen siendo abrumadores. Tan sólo 1/3 de los centros de trabajo españoles (con más de veinte trabajadores) afirman contar con algún tipo de procedimiento para su gestión, concretamente gestionan el estrés un 32.50% de organizaciones, el acoso u hostigamiento un 41% y la violencia física o psíquica un 43.90% (ESENER-2, 2015). Las mismas conclusiones se pueden deducir del informe realizado sobre los riesgos psicosociales en el trabajo por Eurofound y EU-OSHA (2014), donde casi el 80% de los directivos muestran preocupación por el estrés laboral y, sin embargo, apenas han puesto en marcha procedimientos para prevenir estos riesgos. Dicha cuestión se agrava cuanto menor es la plantilla, pues según el informe de ESENER-2 (2015) si un 15% de empresas españolas de más de 20 trabajadores no aplica ninguna medida para prevenir los riesgos psicosociales el porcentaje llega al 46% en las microempresas. Realmente, son las empresas de más de 50 trabajadores las que con más frecuencia aplican planes de acción. Ante estos datos podría sacarse como rápida conclusión que las empresas grandes poseen más cultura preventiva. Sin embargo, la realidad es que este tipo de empresas están sometidas a mecanismos de control más rigurosos en cuanto a prevención se refiere, como por ejemplo la obligación de crear un Comité de Seguridad y Salud (art. 38 LPRL).

Proposición 2:

Son pocas las empresas que cuentan con procedimientos para gestionar riesgos psicosociales tales como estrés, *mobbing*, *burnout* o violencia.

Profundizando en los resultados de la encuesta (ESENER-2, 2015), se señala directamente a la falta de cultura preventiva como principal obstáculo. Así, de las 3 356 empresas españolas encuestadas, el 90.50% reconocieron como principal motivo para gestionar los riesgos psicosociales cumplir la legislación, seguido del 81.10% que aseguró gestionar los riesgos psicosociales para evitar multas y sanciones de la inspección de trabajo. En el conjunto de la Unión Europea se obtuvieron cifras muy similares, 85.40% y 77.90% respectivamente, si bien el número de empresas encuestadas fue muy superior 42 479 (Almodóvar, Pinilla, Galiana y Hervás, 2015). Esta falta de concienciación también se vislumbra a partir del último informe de la ITSS (2015), donde los incumplimientos referidos a la prevención de riesgos psicosociales (52.50%) aumentaron casi un 10% en comparación con el informe del año anterior (43.31%) (ITSS, 2014).

Ante este panorama, no es de extrañar que vayan en aumento los trastornos relacionados con la salud y el bienestar de los trabajadores. Aparte de las campañas de sensibilización sobre la importancia de estos riesgos, es primordial contar con herramientas de evaluación adecuadas que, debido a la falta de criterios claros, deben cumplir con una serie de condiciones que garanticen la calidad de sus mediciones.

2.EVALUACION DE RIESGOS PSICOSOCIALES: INSTRUMENTOS DE MEDIDA

2.1. EVALUACIÓN DE RIESGOS PSICOSOCIALES. MARCO LEGAL

Existen sólidas evidencias científicas que muestran la relación negativa entre la exposición a factores psicosociales de riesgo en el trabajo y la enfermedad. Sabiendo que los riesgos psicosociales prácticamente nunca pueden considerarse inexistentes (Meliá et al., 2006), el empresario debe proteger y garantizar la salud y seguridad de sus trabajadores (art. 14.2 LPRL). Además de ello tiene la obligación de eliminar los riesgos y, en caso de no ser posible, minimizarlos.

Aún sin hacer mención explícita en la LPRL, los riesgos psicosociales deben considerarse incluidos (Bartolomé et al., 2007; Carrasco et al., 2010; Caicoya, 2004; Cruz, De la Flor y Álvarez, 2008; ITSS, 2012; Meliá et al., 2006; Moreno, 2011; Pando, Carrión, Arellano y Saraz, 2006; Peiró, 2010). Dicha inclusión se desprende de la lectura de su articulado (véase tabla 9) y es avalada por numerosa jurisprudencia, tal y como muestra la Sentencia del Tribunal Superior de Justicia de Madrid, Sala de lo Social 2, número 759/2015, donde haciendo referencia al mencionado 14.2 LPRL, se interpretan las situaciones de estrés laboral dentro de la obligación del empresario de proteger y garantizar la salud de sus trabajadores. Por otra parte, en el artículo primero del RSP, sobre integración de la actividad preventiva en las actividades de la empresa, se mencionan específicamente algunos aspectos de la dimensión psicosocial del trabajo como los procesos técnicos, la organización del trabajo y las condiciones del mismo.

Tabla 9: Riesgo psicosocial enmarcado dentro de la Ley de Prevención de Riesgos Laborales
Fuente: Elaboración propia

ARTICULO	TEXTO
4.2	Se entenderá como "riesgo laboral" la posibilidad de que un trabajador sufra un determinado daño derivado del trabajo [...]
4.3	Se considerarán como "daños derivados del trabajo" las enfermedades, patologías o lesiones sufridas con motivo u ocasión del trabajo [...]
4.7	Se entenderá como "condición de trabajo" cualquier característica del mismo que pueda tener una influencia significativa en la generación de riesgos para la seguridad y la salud del trabajador. Quedan específicamente incluidas en esta definición: d) Todas aquellas otras características del trabajo, incluidas las relativas a su organización y ordenación, que influyan en la magnitud de los riesgos a que esté expuesto el trabajador
14.2	En cumplimiento del deber de protección, el empresario deberá garantizar la seguridad y la salud de los trabajadores a su servicio en todos los aspectos relacionados con el trabajo [...]

Continuación tabla 9

15.1	Principios de acción preventiva: d) Adaptar el trabajo a la persona, en particular en lo que respecta a la concepción de los puestos de trabajo, así como a la elección de los equipos y los métodos de trabajo y de producción, con miras, en particular, a atenuar el trabajo monótono y repetitivo y a reducir los efectos del mismo en la salud. g) Planificar la prevención, buscando un conjunto coherente que integre en ella la técnica, la organización del trabajo, las condiciones de trabajo, las relaciones sociales y la influencia de los factores ambientales en el trabajo.
16.2.a	El empresario deberá realizar una evaluación inicial de los riesgos para la seguridad y salud de los trabajadores, teniendo en cuenta, con carácter general, la naturaleza de la actividad, las características de los puestos de trabajo existentes y de los trabajadores que deban desempeñarlos [...]

Para concretar, legalmente se reconocen tanto a la organización como a la ordenación del trabajo como condiciones laborales susceptibles de generar riesgos, y dado que toda empresa consta forzosamente de una organización, éste es un riesgo que no se puede eliminar (Meliá, et al., 2006). De acuerdo con el art. 15 de la LPRL, hay que evaluar los riesgos que no se pueden evitar, por lo tanto, es un imperativo legal evaluar los riesgos psicosociales en todas las organizaciones, lo que implica identificarlos y medirlos, al igual que ocurre con los riesgos higiénicos, de seguridad o ergonómicos. Aparte de ser una imposición legal, Meliá (2008) indica dos razones más para tener en cuenta la evaluación de riesgos psicosociales: (1) las relaciones factores psicosociales-seguridad-salud-comportamientos adictivos (tabaco, alcohol) y (2) la relación factores psicosociales-calidad-productividad de la empresa.

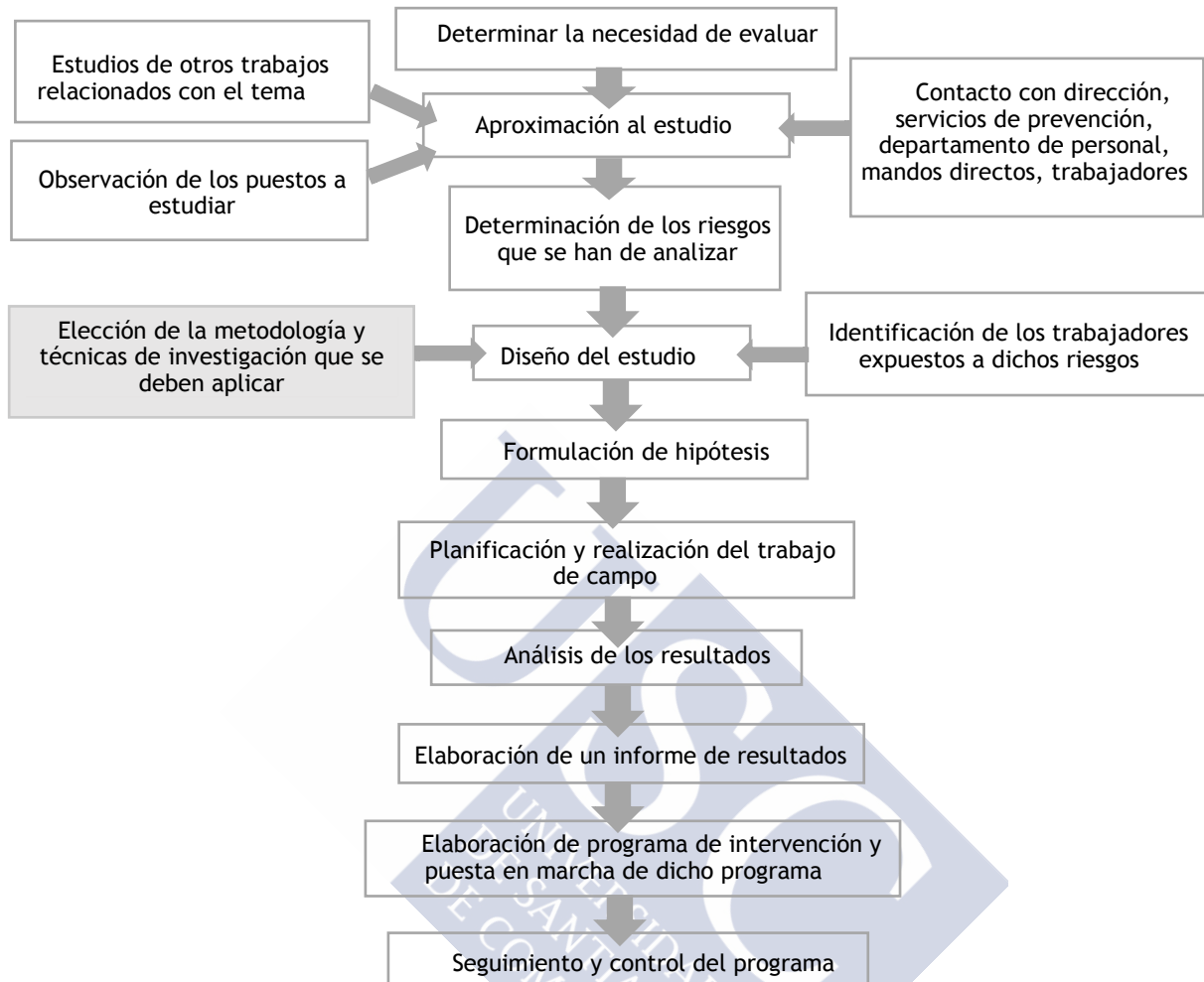
2.2. METODOLOGÍAS Y TÉCNICAS DE EVALUACIÓN DE FACTORES DE ORIGEN PSICOSOCIAL EN EL TRABAJO

La evaluación de riesgos implica estimar la magnitud de los mismos. Para ello es necesario aplicar una metodología que posibilite cuantificar el posible grado de exposición a los factores de riesgo nocivos, a la vez que facilite información al empleador sobre las medidas preventivas a adoptar en cada situación de riesgo (Meliá et al., 2006). De acuerdo con la definición de evaluación de riesgos laborales recogida en el artículo 3.1 del RSP, ésta es:

El proceso dirigido a estimar la magnitud de aquellos riesgos que no hayan podido evitarse, obteniendo la información necesaria para que el empresario esté en condiciones de tomar una decisión apropiada sobre la necesidad de adoptar medidas preventivas y, en tal caso, sobre el tipo de medidas que deben adoptarse.

En vista de que los riesgos psicosociales son intangibles, se puede sentenciar que evaluarlos no es un cometido sencillo. Conforme a la NTP 450 (Oncins y Almodóvar, 1997), la evaluación de riesgos psicosociales es una investigación científica de las condiciones del trabajo donde ha de poder llegarse a un juicio de valor que permita al empresario tomar decisiones enfocadas a garantizar la seguridad y salud del trabajador. La referida nota técnica también apunta que son procesos complejos que constan de varias etapas y además facilita una detallada descripción de cada una de las consecutivas fases que deben llevarse a cabo. En la figura 2 se reproduce el esquema incluido en la NTP 450, el cual resume los pasos necesarios para una adecuada evaluación de los factores psicosociales de riesgo.

Figura 2: Fases evaluación factores psicosociales
(Fuente: Adaptado de NTP 450)



Nota: se ha resaltado en gris la elección de “la metodología y técnicas de investigación que se deben aplicar” por formar parte de los objetivos de esta tesis.

Como resulta obvio tras observar la figura 2, una evaluación de riesgos psicosociales no se fundamenta en el mero hecho de aplicar un simple cuestionario a los trabajadores. Es todo un proceso sistemático compuesto por varias fases. De todas ellas, esta tesis fijará su atención en *la elección de la metodología y técnicas de investigación que se deben aplicar*, por ser ésta la fase donde cobra sentido el uso de los cuestionarios de autoinforme, cuyo análisis conforma uno de los principales aportes de este proyecto de investigación.

Antes de avanzar, es primordial entender que la evaluación de riesgos psicosociales no evalúa el estrés, ni el estado de salud del trabajador, sino que trata de identificar aquellas condiciones de trabajo que llevan implícita la probabilidad de causar un daño, es decir, el riesgo de que suceda. Otra de las cuestiones importantes a tener en cuenta es dilucidar qué es método y qué es técnica, pues a menudo se confunden los términos. Sirva para ello la definición recogida en R. Fernández (2010, p. 167): “método es el camino a seguir para realizar la investigación, mediante una serie de operaciones y reglas prefijadas de antemano, aptas para

alcanzar el resultado propuesto. Técnica no es el camino, sino el arte o manera de recorrerlo”. En la misma dirección, Oncins y Almodóvar (1997), en la NTP 450, realizan la diferenciación de ambos conceptos tratando el método de investigación como la parte teórica y la técnica o instrumento seleccionado para el tratamiento de la información como la parte práctica. Por consiguiente, un método de evaluación de riesgos no es el instrumento que se utiliza para cuantificar los factores de riesgo. Es más que un simple cuestionario, un *check-list* o una entrevista; es un procedimiento sistemático. Para no seguir sustentando el enredo, Guàrdia y Però (2010) apuntan a la necesidad de dejar de referirse a métodos de evaluación de riesgos psicosociales cuando lo que se hace es debatir sobre el rigor científico de las técnicas y los procedimientos propuestos para la recogida de datos.

Una vez dilucidada la posible confusión entre método y técnica, conviene reseñar que desde la perspectiva de la investigación científica tradicionalmente se distinguen dos tipos de enfoques: el cualitativo y el cuantitativo (Garrido y Álvaro, 2007), enfoques que han dado lugar a abundante discusión a la hora de posicionarse por uno u otro (Bueno, 2014; Chavarría, 2011). El debate escapa a los propósitos de este trabajo, aunque sin adentrarnos en él sí conviene conocer las técnicas empleadas en cada uno, dentro de la esfera de la prevención de riesgos psicosociales. Así, según la literatura, nos encontramos con que (1) las técnicas más características del enfoque cualitativo son las entrevistas, grupos de discusión, grupos focales, dinámicas de grupo y registros narrativos; siendo (2) los más populares dentro del enfoque cuantitativo los cuestionarios, las escalas y los *check-list* (Guàrdia y Però, 2010; Meliá et al., 2006; C. Nogareda y Almodóvar, 2006; Oncins y Almodóvar, 1997).

A día de hoy, en España no existen criterios legales específicos (Barreiro, 2006; Cruz, De la Flor y Álvarez, 2008; Ferradans, 2008; ITSS, 2012; Schaufeli y Salanova, 2002). No obstante, destaca el uso que se hace de los cuestionarios y escalas de autoinforme como única técnica de evaluación (Ferrer, 2010), a pesar de ser muchas las voces que coinciden en destacar que lo acertado es combinar ambas. Concretamente, se recomienda utilizar las herramientas cualitativas como complemento de las cuantitativas (Artacoz y Molinero, 2004; Ferrer, 2010; Guàrdia y Però, 2010; ITSS, 2012; Meliá et al., 2006; Moreno y Báez, 2009, 2010; C. Nogareda y Almodóvar, 2006; Oncins y Almodóvar, 1997). Desde la Guía de actuaciones de la Inspección de Trabajo y Seguridad Social sobre riesgos psicosociales (ITSS, 2012), trasladan que la técnica idónea a utilizar dependerá de cada caso en concreto, una vez analizados sus pros y sus contras, por lo que no puede considerarse que haya una técnica mejor que otra. En ese sentido y dependiendo del objeto de análisis, J.M. Cortés (2012) realiza las siguientes recomendaciones (ver tabla 10).

Tabla 10: Técnicas de evaluación de los factores psicosociales dependiendo del objetivo del análisis
Fuente: J.M. Cortés (2012)

HECHOS	OPINIONES	ACTITUDES	COMPORTAMIENTO DEL GRUPO
Observaciones	Encuestas	Entrevistas	Técnicas de grupo
Encuestas	Escalas	Escalas	Observación
Entrevistas	Entrevistas		

Evaluar no sólo se trata de obtener resultados, sino de que esos resultados reflejen de la manera más fiel posible la realidad. Por ello, se hace imprescindible corroborar que las

herramientas seleccionadas cumplan con una serie de condiciones que tienen que ir reflejadas en sus características técnicas, que dirán si son válidas (Moreno y Báez, 2009). Al respecto, Moncada et al. (2014) aseguran que las entrevistas no son adecuadas para la evaluación de riesgos, entre otros motivos porque sus resultados son difíciles de contrastar e interpretar, además de ser necesarios recursos especializados y costosos, a menudo fuera del alcance de la mayoría de empresas. Estas cuestiones han motivado el fomento por parte de las organizaciones líderes a nivel internacional de los cuestionarios estandarizados como “la única estrategia válida y operativa para que la evaluación de riesgos psicosociales pueda ser una realidad en todas las empresas” (p. 15). La citada Guía (ITSS, 2012), indica respecto a la aplicación de técnicas cualitativas que son eficaces en las fases iniciales y posteriores del estudio como forma de recopilación de datos pre-evaluación y para profundizar acerca de los resultados, por lo que se recomienda su utilización para complemento de las técnicas cuantitativas (ej., cuestionarios).

Proposición 3:

A menudo se confunde lo que es un método de evaluación de riesgos (método científico), con el hecho de administrar cuestionarios que evalúan riesgos (técnica de recogida de datos y una de las fases de la metodología

Proposición 4:

La evaluación de riesgos implica más procesos que valorar los riesgos mediante un cuestionario

2.3. CRITERIOS PARA SELECCIONAR LA HERRAMIENTA DE EVALUACIÓN IDÓNEA ATENDIENDO A SU DISEÑO

En la Guía del ITSS (ITSS, 2012) se subraya la importancia de seleccionar procedimientos que proporcionen confianza sobre sus resultados conforme al art. 5 RSP, y en cuanto al diseño del método se marcan los aspectos a tener en cuenta a la hora de la utilización de las técnicas cualitativas y cuantitativas. En lo relativo al uso de técnicas cuantitativas, fundamentalmente basadas en cuestionarios, se tendrá que valorar: i) La fundamentación teórica, ii) la calidad psicométrica del instrumento (fiabilidad y validez), iii) el análisis de los ítems, y iv) la baremación.

La viabilidad de la evaluación de riesgos psicosociales en la empresa se sustenta en el hecho de que se dispone de instrumentos válidos y fiables, con las garantías de calidad necesarias y con fundamento en el conocimiento científico sobre los factores implicados (Meliá et al., 2006, p.24).

Sin duda, la calidad de los procedimientos para la evaluación de riesgos psicosociales va a depender de su fundamentación teórica, pues son necesarios fundamentos científicos sólidos de los factores a evaluar, como de la metodología utilizada para dicha evaluación (Meliá, 2008; Meliá et al., 2006; Peiró, 2010). Los modelos teóricos desarrollados a partir de la investigación científica son de vital importancia a la hora de diseñar métodos de evaluación y prevención válidos, pues de ellos va a depender su capacidad explicativa y predictora de los riesgos psicosociales (Peiró, 2010).

2.3.1. Modelos Teóricos explicativos del estrés.

Muchos autores coinciden al afirmar que la posibilidad de que los factores de riesgo psicosocial afecten negativamente a la salud, se manifiesta a través de las respuestas al estrés, utilizando mecanismos emocionales, cognitivos, conductuales y fisiológicos (Leka y Jain, 2010; Moncada y Llorens, 2006; Moreno y Báez, 2010; Theorell, 1999; 2012). De hecho, la evaluación del estrés ha sido esencial para el desarrollo de la mayoría de los instrumentos de evaluación de riesgos psicosociales en el trabajo (Guàrdia, Però y Barrios, 2008). De este modo, es preciso establecer marcos teóricos que definan y delimiten las características laborales productoras de estrés y poder convertirlas así en datos cuantificables, gracias a los enfoques de investigación tanto cualitativos como cuantitativos (Fernández-López, 2008). La importancia de la fundamentación teórica, en la que se basan las distintas metodologías de la evaluación de riesgos psicosociales, ha sido bien argumentada por palabras de Moreno y Báez (2009, p. 60):

El marco teórico no tiene la finalidad de dar erudición académica, ni tampoco científica, a una propuesta, sino explicabilidad lógica y causal a un conjunto de propuestas. El conocimiento difiere de la información en su orden y organización, y finalmente en su predictibilidad. Desde esta perspectiva, el marco teórico en las metodologías de evaluación de riesgos psicosociales sirve para vincularlo a trabajos anteriores y a resultados previos, en definitiva, a un marco comprobado de conocimientos. La existencia de marcos teóricos no excluye la innovación, sino que la innovación supone la referencia a algún tipo de marco teórico.

Al existir diversidad de fuentes de estrés laboral, existe también diversidad de modelos teóricos o explicativos en los que se basan. Por ejemplo, Kompier (2002) identificó siete teorías principales: (1) Principios socio-técnicos (Cherns, 1976); (2) el Modelo de las características del puesto (MCP) de Hackman y Oldham (1980); (3) Teoría del conflicto de rol expuesta en el modelo de estrés organizacional de Michigan de Kahn, Wolf, Quinn, Snoek y Rosenthal (1964) y en el Modelo de adaptación persona-ambiente de J.R.P. Jr. French, Caplan y Van Harrison (1982); (4) la teoría de la acción de Hacker (Frese y Zapf, 1994; Hacker, Skell y Straub; 1968); (5) Modelo demanda-control-apoyo social de Karasek y Theorell (1990); (6) El modelo vitamínico de Warr (1987, 1994); y (7) Modelo esfuerzo-recompensa de Siegrist (1998).

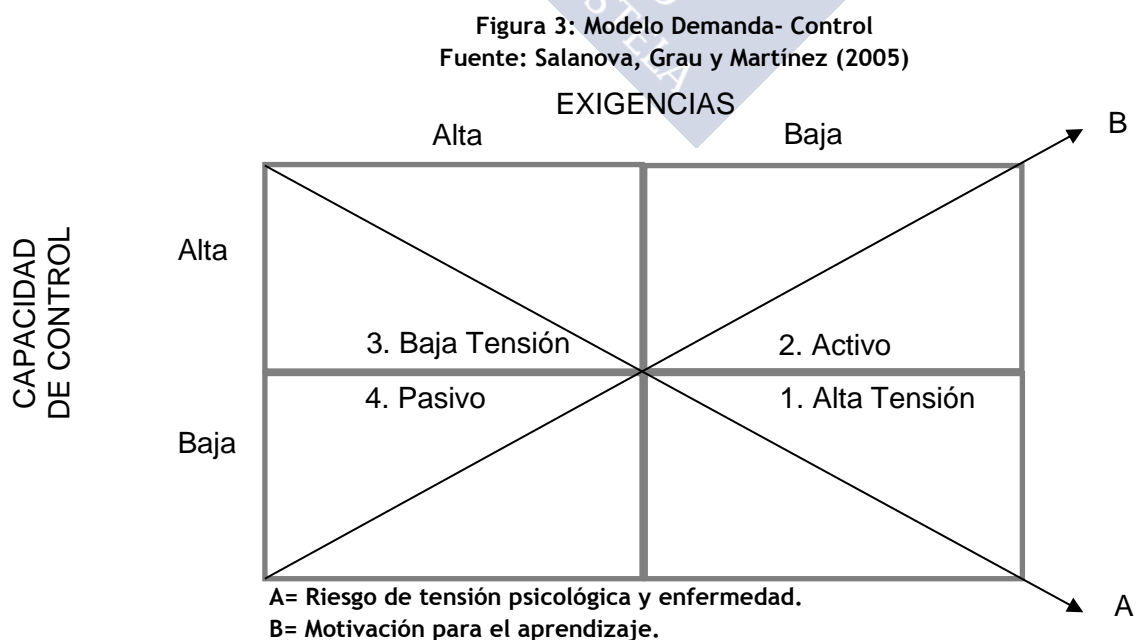
A día de hoy los modelos de investigación sobre el estrés se han ampliado y redefinido. Para muestra las interesantes revisiones de Bucci (2012), Llorens, Líbano y Salanova (2009) o Peiró (2009), entre otras. Sin embargo, para el trabajo que nos ocupa, sólo se hará referencia a los dos modelos que, según la literatura científica, cuentan con mayor evidencia empírica (Barreiro, 2006; Bucci, 2012; Fernández-López, Fernández-Fidalgo y Siegrist, 2005; C. Molina, 2010), a saber: el Modelo Demanda-Control de Robert Karasek (1979) y el Modelo Esfuerzo-Recompensa propuesto por Johannes Siegrist (1996a, 1998). Del estudio de ambos enfoques surgieron sendos instrumentos de evaluación. Por un lado, el Cuestionario de contenidos del trabajo de Karasek et al. (1998) y, por otro, el Cuestionario del desequilibrio esfuerzo-recompensa de Siegrist (1996a).

2.3.1.1. Modelo Demanda-Control

El modelo de *Tensión Laboral* de Karasek (1979), denominado principalmente Modelo Demanda-Control, es uno de los más influyentes en las investigaciones sobre estrés laboral (Serrano M.A., Moya-Albiol y Salvador, 2009 y Vega, 2003). En su modelo original, representado en la figura 3, Karasek (1979) propone una relación bidimensional entre la demanda del trabajo (entendida como las exigencias laborales tanto cualitativas como cuantitativas) y el control con el que cuenta el trabajador para responder a esas demandas. Esta última variable está compuesta por dos sub-escalas: la autonomía y la capacidad de desarrollar habilidades propias (Karasek y Theorell, 1990).

La relación entre las exigencias del trabajo y la dimensión positiva del control actuará como predictor de situaciones de tensión emocional (diagonal A de la figura 3) y de comportamiento activo/pasivo (diagonal B de la figura 3). La dimensión fundamental es el control por ser capaz de moderar las exigencias del trabajo, es decir, “el estrés no depende tanto del hecho de tener muchas demandas, como de no tener capacidad de control para resolverlas” (Vega, 2003, p. 2).

De la combinación de ambas variables y fuentes de estrés, resultan cuatro posibles situaciones: (1) mucha tensión: este nivel está integrado por altas exigencias y poco control para afrontarlas, condición generadora de estrés y consecuentemente, mayor predisposición a padecer enfermedades cardiovasculares, trastornos musculo-esqueléticos, ansiedad, alergias, hipertensión, procesos asmáticos, entre otras dolencias (Karasek y Theorell, 1990; Schnall, et al., 2000); (2) trabajo activo: caracterizado por presentar elevadas demandas y alta capacidad de decisión, creando la condición de desafío. Este nivel aumenta la motivación y las oportunidades de uso y aprendizaje de habilidades, amortiguando los efectos del estrés negativo; (3) poca tensión: en este cuadrante confluyen las bajas demandas con el alto control; según Vega (2003) sería un estado parecido al de relajación; (4) trabajo pasivo, consistente en mínimas exigencias y bajo control, empleos poco interesantes donde es fácil encontrarse en la situación de falta de motivación y pérdida paulatina de las habilidades ya aprendidas, dando lugar al aprendizaje negativo.

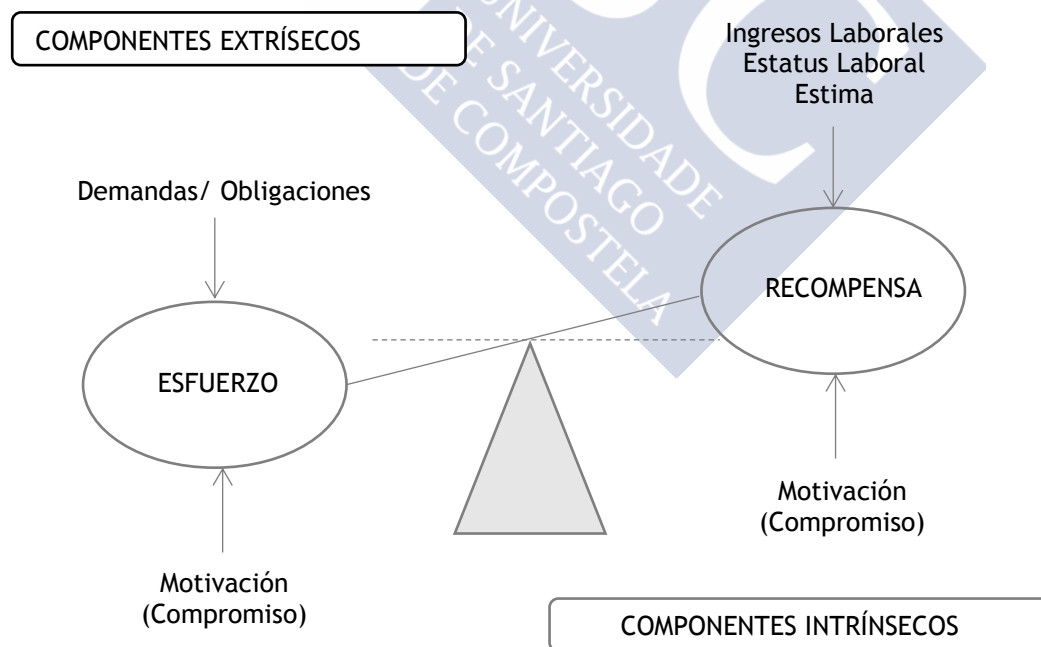


Johnson y Hall (1988) establecieron el Modelo Demanda, Control y Apoyo social al añadir una tercera variable, la dimensión del apoyo social en el trabajo, que se refiere no sólo a la cantidad de relaciones sociales sino también al grado del apoyo prestado por compañeros y superiores. Esta variable tendría una doble vertiente. Cuando es positiva, moderaría los efectos del estrés ocasionado por altas demandas y bajo control, sin embargo, cuando es bajo o inexistente se convertiría en un factor de estrés independiente, un estresor más a añadir a los ya existentes. Karasek y Theorell, (1990) en la reformulación de su modelo, sugieren que bajo condiciones de altas demandas, bajo control y escaso apoyo social el trabajador podría encontrarse en situación de aislamiento y alta tensión.

2.3.1.2. Modelo Esfuerzo-Recompensa (ER)

Propuesto por Siegrist (1996a, 1998), refleja la influencia del rol laboral en las personas, explicando el estrés laboral como consecuencia de un desequilibrio entre el esfuerzo que realiza el trabajador y las recompensas recibidas. Las variables de esfuerzo de este modelo distinguen entre las exigencias extrínsecas, referidas a las cargas laborales y obligaciones; y las exigencias intrínsecas, que describen las características personales de motivación y afrontamiento en relación con las demandas del trabajo. En cuanto a las recompensas, éstas no abarcarían sólo las ganancias en términos salariales, sino que incluyen el apoyo social y estima por parte de compañeros y superiores, así como también se refieren a las oportunidades profesionales y la seguridad en el trabajo (ver figura 4).

Figura 4: Modelo Esfuerzo-Recompensa
(Fuente: Siegrist, 1996a).



Relacionando las variables anteriores, Siegrist (1996a, 1998) y Siegrist, Peter, Cremer y Seidel (1997) proponen que un alto esfuerzo aparejado a una baja recompensa, mantenido en el tiempo, propiciaría el aumento de tensión, constituyendo la situación de mayor riesgo para la salud. Así puede ratificarse en Fernández-López et al. (2005), donde se muestra un resumen de

distintas evidencias empíricas de predicción del modelo ER, hallándose asociaciones significativas de riesgo cardiovascular, hipertensión, dependencia alcohólica, depresión, alteraciones psiquiátricas, colesterol y otras.

Estos modelos lejos de ser antagónicos, son modelos complementarios (Luceño, Martín, Jaén y Díaz, 2005; Peña, Ramos y Martínez-Losa, 2009; Siegrist, 2002) y han sido utilizados para la formulación de diferentes instrumentos de evaluación de riesgos psicosociales, como los cuestionarios de auto-informe del método ISTAS 21, DECORE, INSL y el método WONT (Giner, 2012; Lahera, 2006; Lahera y Nogareda, 2009; Luceño et al., 2005; Moncada, Llorens, Kristesen y Vega, 2005; Peiró et al., 2013).

Aunque en los últimos años hayan avanzado sustancialmente la formulación de teorías sobre el estrés y su entendimiento como factor de riesgo para la salud, los estresores siguen siendo muy difíciles de delimitar y cuantificar, y he aquí que nos topamos con la compleja tarea de medir lo que no se ve.

2.3.2. Medición de los factores psicosociales de riesgo

Para Meliá (2008, p. 3) “sin lugar a dudas la evaluación de los riesgos psicosociales es un tipo de evaluación psicológica aplicada, enraizada con la psicología de la seguridad y salud ocupacional y con la psicometría como ciencia técnica de la medición psicológica”. En la misma tesitura se sitúan Guàrdia y Però (2010) cuando afirman que la psicopsicología utiliza la métrica psicológica. Ésta, para poder predecir, evaluar o explicar los comportamientos de las personas, tiene que medir los rasgos y atributos psicológicos del ser humano, características que a menudo están compuestas de variables complejas e inobservables -y por tanto de difícil definición-, como son la inteligencia, la inseguridad, la autoestima, la depresión, el estrés o la creatividad; fenómenos también denominados *constructos*, que sólo son posibles de medir a través de la observación de las conductas a las que dan lugar (Aliaga, 2006; Cortada de Kohan, 2001; Crocker y Algina, 1986; Cronbach y Meehl, 1995; Gregory, 2012; Nunnally, 1978). De este modo, cuando hablamos de cuantificar dimensiones psicosociales tales como inseguridad, estima, apoyo social o liderazgo, estamos también ante otros constructos psicológicos. Al igual que acontece con la medición de lo *psicológico*, los riesgos psicosociales llevan implícita una característica notoria que dificulta su evaluación: su intangibilidad.

En relación a la medida de constructos, Thurstone (1928) afirmó que las actitudes pueden medirse. Sin embargo, las peculiares características inherentes de las variables psicológicas hacen que su medición resulte compleja y que no se dispongan de instrumentos para su medición directa, pues su naturaleza intangible impide que sean directamente manipulables, como lo pueden ser otro tipo de variables propias de las ciencias físicas (Abad, Garrido, Olea y Ponsoda, 2006; Muñiz, 1998a, 2005; Viladrich y Doval, 2011). A modo ilustrativo, cabe apuntar que medir la cantidad de agua que contiene un recipiente dado no conlleva la misma complejidad que medir por ejemplo la cantidad de inteligencia, autoestima, estrés o seguridad, que *contiene* una persona. Para solventar este obstáculo y poder realizar mediciones de los procesos psicológicos con garantías de rigor científico, surgió la *Psicometría*, que se ocupa de cuantificar y dar formato a dichas mediciones, utilizando el lenguaje de las matemáticas (Aliaga, 2006; Jáñez, 1989; Meliá, 1990).

Como nota común entra la multitud de definiciones existentes, se puede admitir que la psicometría se ocupa de los problemas de medición de las variables psicológicas (Aliaga, 2006; Cliff, 1979; Maciá, 1984; Muñiz, 1998a; Nunnally y Bernstein, 1995; Però y Guàrdia, 2010; F. Rivas, 1979; Yela, 1968, 1987). Entre sus objetivos está cuantificar dichos fenómenos, es decir, transformar los hechos en datos (Cliff, 1979), diseñando instrumentos adecuados que permitan realizar inferencias, a través de la conducta, y así poder cuantificar las diferencias individuales (Bunge, 1973; Muñiz, 2005; 2010).

Entre estas técnicas se encuentran el escalamiento de los sujetos, que no son otra cosa que test psicométricos (Muñiz, 1998a), los cuales están fuertemente arraigados en la cultura occidental, pues incluyen un procedimiento de puntuación con los que se obtienen medidas que son utilizadas para muy diversos fines, como selección de personal para un puesto de trabajo, acceso a programas educativos, diagnósticos clínicos, o estimación el grado de las personas en un constructo concreto como, por ejemplo, ansiedad, estrés, depresión, seguridad, autonomía, entre otros (Delgado, 1998; Prieto y Delgado, 2010).

En consecuencia, resulta difícil obtener mediciones de manera objetiva teniendo que acceder a los resultados a través de técnicas denominadas *subjetivas* como son la observación, entrevistas a trabajadores o los test, entre los que se encuentran los cuestionarios de autoinforme. Dado que estos últimos son una de las herramientas más utilizadas para la evaluación de los riesgos psicosociales (Artazcoz y Molinero, 2004; Blanch, Sahagún y Cervantes, 2010; Caicoya, 2004; Ferradans, 2008; Ferrer, 2010; Gimeno, 2004; Giner, 2012; Guàrdia, 2010a, 2010b; Guàrdia y Però, 2010; ITSS, 2012; Leka y Jain, 2010; Luceño; Marín, Díaz et al., 2008; C. Molina y García, 2010; Moreno y Báez, 2010; Moorman y Podsakoff, 1992; Peiró, 2010; Rick, Briner, Daniels, Perryman y Guppy, 2001), conviene realizar un análisis más profundo sobre los test psicométricos y sus garantías científicas a la hora de su construcción y utilización.

2.3.3. Test Psicométricos: Conceptualización y clasificación

Los test psicométricos, aunque se basan en información de naturaleza subjetiva, son cuestionarios utilizados como instrumentos de medida objetiva y tipificada de los constructos o variables psicológicas (Anastasi, 1988; Aliaga, 2006; Viladrich y Doval, 2011), entre las que se encuentran las variables relacionadas con los riesgos psicosociales. Una completa definición de que lo que es un test la expone Muñiz (2005, p. 9): “empezando por lo obvio, hay que recordar que las respuestas a los test son conductas, pudiendo definirse un test como una muestra de conducta de una persona recogida de forma objetiva y estandarizada en un momento determinado”. Los test psicométricos implican medidas lo más libres posibles de interpretaciones y comparables con otras medidas de sujetos suficientemente representativas (E.M. García y Magaz, 2009).

Desde el nacimiento de los test en el laboratorio de Galton en 1884 (citado en Buela-Casal y Sierra, 1997) hasta nuestros días, los test han experimentado una notable expansión, existiendo infinidad de ellos (se recomienda la lectura de los trabajos publicados de Anastasi, 1998; Muñiz, 1998a y Navas, 1999, para un mejor conocimiento de la historia y evolución de los test psicométricos). Son utilizados en áreas muy diversas como Psicología, Sociología, Psicosociología o Educación, midiendo desde los procesos más básicos hasta aspectos más

complejos, como rasgos de personalidad o aspectos cognoscitivos, sin olvidarnos de los test que miden actitudes o valores encuadrados en el ámbito sociocultural (Muñiz, 1998a). En palabras de Muñiz, (2010, p. 58) “la historia de los test es una historia exitosa de la que la psicología tiene que sentirse orgullosa”.

En lo concerniente a los tipos de test, son tan diversos y numerosos que se hace necesario ordenarlos por categorías. Obviamente, la gran diversidad de test implica otros tantos modos posibles de clasificarlos, existiendo numerosa literatura al respecto. Con el propósito de intentar esclarecer de una manera asequible los principales tipos de test, se incluye en la tabla 11 la clasificación ofrecida por Aliaga (2006, p. 89), por ser sencilla y comprensible.

Tabla 11: Clasificación test psicométricos
Fuente: Adaptación realizada a partir de Aliaga (2006)

CRITERIO	CLASIFICACIÓN
Por su forma de dar las instrucciones	Orales Escritos
Por su administración	Individual Colectiva Auto-administrados
Por la forma o tipo de respuesta que exigen	Objetivas Subjetivas
Por el material de la prueba	Lápiz y papel Verbal completamente De ejecución (material, manual, viso motor) De aparatos especiales Combinación de los tres primeros (Ej.: WAIS)
Por su forma de calificación	Manual Electrónica
Por sus edades límites de aplicación	Pruebas para infantes (baby test) Pruebas para preescolares Pruebas para escolares Pruebas para adolescentes Pruebas para adultos
Por su libertad de ejecución	Pruebas de poder (power test) Pruebas de rapidez (speedy test)

Siguiendo esta clasificación, y en lo que concierne a los test psicométricos más utilizados para evaluar riesgos psicosociales, éstos suelen ser: escritos, individuales, de respuesta subjetiva, de lápiz y papel y pueden calificarse de manera manual y/o electrónica.

2.3.4. Rigor científico en los Test

A pesar de su diversidad de aplicaciones, no todo ha sido un camino fácil en la historia de los test, pues han surgido multitud de posturas críticas sobre su uso (Fernández-Ballesteros, 2013). Concretamente, en la década de los 70 fueron objeto de profunda discusión social y litigio por considerarlos injustos y sesgados (Delgado, 1998, Navas, 1999). La crítica ha sido especialmente feroz con la evaluación psicológica realizada mediante test con fines de selección tanto educativos como laborales, preservándose de la quema los test utilizados con motivo de evaluación diagnóstica, sobre todo en el ámbito de la salud (Navas, 1999). Para salvar estas

críticas es preciso una buena utilización de los test, cerciorarse de que son justos, evitando los sesgos, así como garantizar sus propiedades psicométricas (Gómez-Benito, Hidalgo y Guilera, 2010; Muñiz, 2010; Ponsoda, 2009; Prieto y Delgado, 2010).

Como se ha mencionado en párrafos precedentes, los test son instrumentos de medida que sirven para realizar inferencias. Para corroborar que esas inferencias sean las adecuadas es preciso aplicar las teorías estadísticas de los test, ya que sin estas teorías no se podrían utilizar con garantías de rigor científico. Así, en cuanto a medición psicométrica se refiere, es imprescindible que las mediciones realizadas sean precisas y contengan el mínimo error posible (Muñiz, 1998a; 2010). Concretamente, es “La Teoría de los Test”, que construye y desarrolla los test psicométricos como herramientas de medida, la que se ocupa de estimar el grado de error derivado de los problemas de medición propios de las variables psicológicas y de confeccionar instrumentos donde se minimicen tales errores (Martínez-Arias, 1996).

Para ello, la Teoría de los Test, sirviéndose de modelos matemáticos y disciplinas como la estadística y el diseño experimental, se ocupará de analizar si el test mide lo que dice medir y el grado de precisión y congruencia de sus registros. En otras palabras, se comprueba si las mediciones realizadas con los test son “fiabiles” y “válidas” (Kerlinger y Lee, 2002). La fiabilidad y la validez son las propiedades psicométricas que aledañas a la fundamentación teórica, permitirán garantizar que las inferencias o predicciones a realizar serán las más adecuadas y eficientes, y no producto de una medición banal (Argibay, 2006; Muñiz, 1998b, 2005, 2010; Prieto y Delgado, 2010).

Con relación al objetivo de esta tesis, cabe decir que la mayoría de los test de autoinforme utilizados en las evaluaciones de riesgos psicosociales como el FPSICO, ISTAS 21 o MC-UB están contruidos bajo la óptica de la Teoría Clásica de los Test (Peró y Guàrdia, 2010) que, a su vez, se basan en el modelo matemático propuesto por Spearman (1904, 1907, 1913, citados en Muñiz, 2010) y:

Consiste en asumir que la puntuación que una persona obtiene en un test, que denominamos su puntuación empírica, y que suele designarse con la letra X , está formada por dos componentes, por un lado la puntuación verdadera de esa persona en ese test (V), sea la que sea, y por otro un error (e), que puede ser debido a muchas causas que se nos escapan y que no controlamos. Lo dicho puede expresarse formalmente así: $X = V + e$ (Muñiz, 2010, p. 60).

De lo citado anteriormente se desprende que los instrumentos utilizados serán más fiables cuanto menor sea el error, y esto se consigue cuando el valor verdadero (X) se encuentre lo más próximo posible al valor observado. Cuanto mayor es el error, más contaminado estaría el test; por el contrario, al minimizar el error se incrementan la fiabilidad y la validez del instrumento utilizado (Argibay, 2006).

2.3.4.1. Fiabilidad

La fiabilidad se refiere a la precisión, estabilidad o grado de consistencia del instrumento utilizado; es decir, en qué medida el test utilizado es preciso al medir los valores verdaderos de los constructos analizados cuando el proceso de medición se realiza repetidas veces. La fiabilidad está directamente relacionada con los errores de medición, pues un test será más

fiable cuanto menor sea el error, es decir, cuanto más logre aproximar el valor observado al valor verdadero (Abad et al, 2006; Argibay, 2006; León y Montero, 2003; Prieto y Delgado, 2010).

Como se asume que la puntuación verdadera de las personas debe ser siempre la misma cada vez que se repita el test (dando por hecho que la variable a medir se mantiene constante en el tiempo), si existe alguna variación en sus resultados es debido al error aleatorio, que es un error impredecible e incontrolable, ocupándose de este tipo de error el examen de la fiabilidad (Argibay, 2006).

Existen varios modos de estimar el grado de confiabilidad de un test: (i) fiabilidad test-retest, (ii) fiabilidad o coeficiente consistencia interna, (iii) método de formas alternativas o paralelas, (iv) método interjueces o interevaluadores y (v) método de mitades partidas. Todas estas técnicas utilizan fórmulas que proporcionan coeficientes de fiabilidad (R. Hernández, Fernández y Baptista, 2014; Moreno-Rosset, 2005; Prieto y Delgado, 2010) muy útiles para optar por el test más preciso entre todos los posibles, pues a mayor fiabilidad mayor correspondencia entre los valores observados y los verdaderos (Argibay, 2006; Martínez-Arias, 1995).

2.3.4.2. Validez

Una condición indispensable que debe reunir un test se refiere a su índice de validez; es decir, en qué grado mide lo que pretende medir (Gregory, 2012; Sanz, Izquierdo y García-Vera; 2013). Sin embargo, que las mediciones de un test sean fiables no implica que sean válidas. Gregory (2012) afirma que “los psicólogos reconocen desde hace mucho que la validez es la característica fundamental y más importante de una prueba ya que, después de todo, es lo que define el significado de las puntuaciones obtenidas” (p. 110). Muñiz (1998b), va más allá, diciendo que la validez es la “clave de la medición en ciencias sociales” (p. 13.)

A diferencia de la fiabilidad, a la cual interesa el error aleatorio, a la validez le interesa el error sistemático (Aliaga, 2006). En relación con la definición de validez, los estándares para los test educativos y psicológicos se refieren a ella como el grado de adecuación, significación y utilidad de las inferencias específicas realizadas a partir de las puntuaciones obtenidas de un test (*Standards for Educational and Psychological Testing*, AERA, APA y NCME, 2014). Así, un test con un índice óptimo de validez permitirá que las predicciones que se realicen partiendo de sus resultados sean correctas. Por lo tanto, lo que se valida no es el test, sino las inferencias realizadas a partir de él (Aliaga, 2006; Cronbach, 1971; Moreno-Rosset, 2005; Muñiz, 1998b, 2005; Silva, 1989). Parafraseando a Muñiz (1998b), “la forma estándar de validar las inferencias es derivar predicciones y contrastarlas con los datos. Con sus luces y sus sombras, el método hipotético deductivo experimental sigue siendo el canon para la validación” (p. 13).

APA, AERA y NCM (2014), que representan el consenso en la comunidad psicométrica, indican que la validez es el proceso de acumulación de evidencia empírica en apoyo a tales inferencias o hipótesis. Aunque existen diversos procedimientos de acumular estas evidencias, tradicionalmente se han agrupado en tres categorías: validez de contenido, validez de constructo y validez de criterio (Abad et al., 2006; Aliaga, 2006; Argibay, 2006, E. M. García y Magaz, 2009; Gregory, 2012; R. Hernández et al., 2014; Muñiz, 1998b; Prieto y Delgado, 2010).

Dependiendo del propósito, interesará conocer uno u otro tipo de validez. A continuación, se recoge una breve descripción de las tres condiciones:

2.3.4.2.1. Validez de contenido

Trata de comprobar que la muestra de ítems utilizada en la elaboración del test es significativa y representativa del universo posible del constructo a medir (Argibay, 2006). Suele estar sustentada en análisis lógicos del contenido de los ítems y expresarse cualitativamente, por lo que su comprobación no se realiza mediante procedimientos estadísticos, sino que se consulta a *jueces* o expertos, preferiblemente en número impar, acerca de su grado de acuerdo sobre la representatividad de los reactivos en el dominio de interés (Aliaga, 2006).

Lo deseable, aunque no siempre sucede así, es que los autores del test hagan explícita la definición operativa de la variable o variables que mide su test. Si esta definición no se encuentra de manera explícita en su manual técnico, entonces debemos considerar si, a nuestro juicio, los elementos del test corresponden realmente a manifestaciones de la variable que el profesional desea medir (E.M. García y Magaz, 2009, p. 21).

2.3.4.2.2. Validez de constructo

La validez de constructo se refiere a la obtención de evidencias que corroboran que las conductas registradas en el test son indicadores válidos del constructo que se pretende medir. Dicho de otro modo, es el grado en que los ítems miden la variable a la que se refieren (Aliaga, 2006; Argibay, 2006). Las formas de recoger estas evidencias suelen estar enfocadas en el Análisis Multivariado, siendo muy habitual la utilización de técnicas de Análisis Factorial tanto exploratorio como confirmatorio (Abad et al., 2006).

2.3.4.2.3. Validez de criterio o predictiva

Este tipo de validez corresponderá estudiarla si el propósito del test es realzar predicciones, ya que las evidencias sobre la validez de criterio lo que muestran es el grado en que las puntuaciones de un test predicen o pronostican variables de interés que actúan como estándar.

Dentro de la validez de criterio se distingue entre (i) validez predictiva y (ii) validez concurrente. La distinción entre ambas reside en el posicionamiento la temporal del criterio, así la validez predictiva indica el grado de eficacia de la prueba en el pronóstico de algún resultado futuro, y la validez concurrente sería el grado en que la puntuación de una prueba se relaciona con alguna medida criterio tomada al mismo tiempo (González-Llaneza, 2007; Moreno-Rosset, 2005). La obtención de estas evidencias, tanto predictiva como concurrente, se realiza mediante un coeficiente de correlación entre las puntuaciones del test y las variables utilizadas como criterio (Abad, et al., 2006; Aliaga, 2006).

2.3.5. Cuestionario de autoinforme aplicados a la evaluación de riesgos psicosociales.

Dentro de la categoría de test psicométricos se encuentran los cuestionarios de autoinforme o inventario autoadministrado. Kent (2003, p. 425) los define del siguiente modo: “cuestionario donde los entrevistados responden preguntas sobre sí mismos”. En cuanto a este propósito, los autoinformes son útiles para obtener información que no es posible conseguir mediante la simple observación, pues pueden referirse a multitud de manifestaciones tanto internas como

externas, entre ellas conductas motoras, respuestas fisiológicas, pensamientos, sentimientos, experiencia subjetiva o expectativas futuras (Fernández-Ballesteros, 2013).

De lo que no cabe duda es de que constituyen la técnica de evaluación más utilizada para la recolección de datos, tanto en Psicología como en las Ciencias Sociales (Abad, Olea, Ponsoda y García, 2011; Barker et al., 2002; Fernández-Ballesteros, 2003; Gómez-Benito et al., 2010). Por eso no sorprende que sea a su vez el instrumento más utilizado para la evaluación de los riesgos psicosociales (Artazcoz y Molinero, 2004; Blanch et al., 2010; Caicoya, 2004; Ferradans, 2008; Ferrer, 2010; Gimeno, 2004; Giner, 2012; Guàrdia, 2010a; Guàrdia y Però, 2010; ITSS, 2012; Jensen, 1995; Leka y Jain, 2010; Luceño, Martín, Díaz et al., 2008; C. Molina y García, 2010; Moreno y Báez, 2010; Moorman y Podsakoff, 1992; Peiró, 2010; Rick et al., 2001).

Existen variedad de razones que hacen recomendable el uso de los cuestionarios de autoinforme, entre ellos la relación coste-eficacia, el ahorro de tiempo, la fácil aplicación e interpretación o que permiten contrastar diferencias individuales (De las Cuevas y González de Rivera, 1992; Echeburúa et al., 2003; Fernández-Montalvo y Echeburúa, 2006; Ferrer, 2010; Jensen, 1995; ITSS, 2012; Quick, 1997; Luceño, Martín, Díaz et al., 2008; Moorman y Podsakoff, 1992). Sin embargo, el principal argumento reside en que están contruidos bajo los principios de la psicometría, siendo los únicos que hasta el momento proporcionan datos sobre su fiabilidad y validez (Caicoya, 2004; Luceño, Martín, Díaz et al., 2008).

Con todo, Derogatis y Melisaratos (1983, citado en De las Cuevas y González de Rivera, 1992, p. 1) defienden que la principal ventaja de los cuestionarios autoinformados es su posibilidad de acceso a información exclusiva que procede “desde el mismo que la experimenta”. En la misma línea se expresan Barker et al. (2002):

La gran ventaja del autoinforme es que proporciona directamente los propios puntos de vista de los encuestados. Da acceso a la información fenomenológica, es decir, la percepción de los encuestados de sí mismos y de su mundo, imposibles de conseguir de ningún otro modo (p. 95).

En relación con la evaluación de riesgos psicosociales, hay que tener en cuenta que la experiencia del estrés y sus causas son difíciles de valorar si no se tiene en consideración la propia percepción individual del trabajador que los sufre (Moreno y Báez, 2010; Peiró, 2008; Villalobos, 2004). Según la regla de oro de Kelly (1955), la mejor manera de saber lo que le ocurre a una persona es preguntarle. Cumpliendo esa premisa se encuentran los cuestionarios de evaluación de riesgos psicosociales, cuyas medidas explícitas de exposición están basadas en la autopercepción que el trabajador tiene sobre su propia experiencia laboral (Moorman y Podsakoff, 1992).

Con este tipo de test es posible recoger información objetiva de los trabajadores que cubren el cuestionario. Pero para ello es necesario que las herramientas utilizadas ofrezcan garantías de uso, corroborando el cumplimiento de ciertas propiedades (Guàrdia, 2010a, 2010b; ITSS, 2012; Meliá, 2008). Las más habituales según estos expertos son:

- a. Fundamentación científica.
- b. Contrastes y pruebas específicas del método.

- c. Fiabilidad y validez (incluyendo ausencia de sesgo, Abad et al., 2006; 2011).
- d. Análisis de los ítems.
- e. Baremación, estandarización y administración.
- f. Evaluación de riesgos psicosociales y propuestas de intervención.
- g. Adaptación a la legislación vigente.
- h. Aplicabilidad en la empresa y uso profesional de los instrumentos de evaluación.
- i. Conocimientos del evaluador y pruebas piloto.

De todas las propuestas desarrolladas para evaluar la esfera psicosocial en el entorno laboral, en el presente trabajo sólo serán objeto de mención las que afirmen cumplir con la mayoría de dichos criterios.

Proposición 5:

Para valorar los riesgos psicosociales es esencial la autopercepción del trabajador

Proposición 6:

Para que los datos recogidos sobre la autopercepción del trabajador sobre su experiencia y exposición a los riesgos psicosociales sea lo más objetiva posible, la herramienta a utilizar para la recogida de datos debería contener los siguientes criterios:

- a. Que cumpla propiedades psicométricas de fiabilidad y validez
- b. Que controlen los sesgos
- c. Baremación y estandarización
- d. La fundamentación teórica
- e. Otros criterios

2.4. REVISIÓN DE LAS PRINCIPALES TÉCNICAS BASADAS EN ESCALAS AUTOINFORME UTILIZADAS EN LA EVALUACIÓN DE RIESGOS PSICOSOCIALES

Hoy en día convive una gran diversidad de instrumentos para la investigación de los factores psicosociales de riesgos derivados de las condiciones laborales. Blanch et al. (2010) los dividen en tres categorías: (i) los que aportan información general acerca de las condiciones del trabajo; (ii) los que investigan sobre un conjunto de factores psicosociales, basados mayormente en cuestionarios de autoinforme; y (iii) herramientas especializadas en informar sobre un determinado factor psicosocial (*burnout*, *mobbing*), o sobre el efecto psicológico derivado del mismo, como ansiedad, depresión, estrés, salud mental. Este tipo de herramientas suelen aplicarse como complemento de los instrumentos dedicados a informar sobre conjuntos de factores, y a menudo sobrepasan el contexto estrictamente laboral, teniendo una función más reactiva que preventiva.

Partiendo de esta categorización, y con objeto de otorgar cierto orden dentro de la gran variedad de técnicas utilizables, en la tabla 12 se expone una clasificación de los instrumentos más populares según la revisión de la literatura.

Tabla 12. Instrumentos de investigación de riesgos psicosociales según información aportada
Fuente: Elaboración propia

1 - INFORMAN SOBRE CONDICIONES DE TRABAJO EN GENERAL
LEST (Guélaud, Beauchessne, Gautrat y Roustang, 1978)
ANACT (INSHT, 1991)
Evaluación de las condiciones de trabajo en PYMES (Bestraten et al., 2000)
Cuestionario Encuesta de Calidad de Vida en el Trabajo (Ministerio de Trabajo y Asuntos Sociales, 2010)
<i>Questionnaire for Fifth European Survey on Working Conditions</i> (Eurofound, 2010)
2 - INFORMAN SOBRE EL CONJUNTO DE FACTORES PSICOSOCIALES
Metodología Prevenlab-Psicosocial de la Universidad de Valencia (Peiró, 1999b)
<i>Perceived Work Characteristics Survey</i> (Haynes, Wall, Bolden y Ricket, 1999)
El cuestionario “FPSICO” (Martín-Daza y Pérez, 1997)
<i>General Nordic Questionnaire</i> (QPS Nordic, Lindstrom, et al., 2000)
Cuestionario de Evaluación de Riesgos Psicosociales, INSL (Lahera y Góngora, 2002)
Método ISTAS-21 (CoPsoQ, Moncada et al., 2002)
INSL (Lahera y Góngora, 2002)
INERMAP (Gracia Camón, 2003)
Manual para la evaluación de riesgos ergonómicos y psicosociales en la PYME (INSHT e IBV, 2003)
Método MARC-UV: BFPSL (Meliá, 2003a) + BATERIA VALENCIANA PREVACC (Meliá, 2003b)
<i>HSE Indicator Tool</i> (Cousins et al., 2004)
Cuestionario multidimensional DECORE (Luceño, Martín, Tobal y Jaén, 2005)
Cuestionario <i>Mini Psychosocial Factors</i> (MPF, E. Ruiz e Idoate, 2005)
Metodología RED-WONT (Salanova, Llorens, Cifre y Martínez, 2006)
Batería MC MUTUAL-UB (Guàrdia, Però y Barrios, 2008)
<i>Work-Related Quality of Life scale</i> (WRQoWL, Van Laar, Edwards y Easton, 2007)
<i>Copenhagen Psychosocial Questionnaires</i> (Pejtersen, Kristensen, Borg y Bjorner, 2010)
<i>Workplace Stressors Assessment Questionnaire</i> (Mahmood, Coons, Guy y Pelletier, 2010)
El QVILA (Blanch et al., 2010; Blanch, Sahagún, Cantera y Cervantes, 2010)
3 - INFORMEN SOBRE LOS FACTORES ESPECÍFICOS
a. ESTRÉS
<i>Job Diagnostic Survey</i> (JDS, Hackman y Oldham 1975)
<i>Ways of Coping</i> (WOC, Folkman y Lazarus, 1980)
<i>Stress Diagnostic Survey</i> (SDS, Ivancevich y Matteson, 1983)
<i>Job Stress Scale</i> (Parker y DeCotiis, 1983)
SCOPE (Brenghelmann, 1986)
<i>Stress at Work</i> (Frese, 1985)
<i>NIOSH Generis Job Stress Questionnaire</i> (Hurrell y McInaney 1988)
<i>Occupational Stress Indicator</i> (OSI, Cooper, Sloan y Williams, 1988)
Cuestionario de estrés ocupacional para profesionales de la salud (Reig y Caruana, 1990)
Índice de reactividad al estrés (IRE-32) (González de Rivera, 1990)
<i>The Stress profile</i> (Wheatley, 1990)
Escala de Apercepción del Estrés (Fernández- Seara, 1992)
<i>Working Conditions and Control Questionnaire</i> (WOCCQ, De Keyser and Hansez 1996)
VBBA (Van Veldhoven, Meijman, Broersen y Fortuin, 1997)
<i>Occupational Stress Inventory</i> (OSI-R, Osipow, 1998)
<i>Job Content Questionnaire</i> (JCQ, Karasek et al., 1998)
<i>Pressure Management Indicator</i> (A. Williams y Cooper, 1998)

Continuación tabla 12

Job Stress Survey (JSS, Vagg and Spielberg 1999)

Escala de fuentes de estrés en profesores (EFEP, S. Nogareda, 2001)

Job Stress (Lambert, Hogan, Camp y Ventura, 2006)

Stress in General (versión revisada, Yankelevich, Broadfoot, Gillespie, Gillespie y Guidroz, 2012)

b. BURNOUT

Staff Burnout Scale (J.W. Jones, 1980)

Indicadores del Burnout (Gillespie, 1980)

Tedium Measure (Pines, Aronson y Kafry, 1981)

Maslach Burnout Inventory (MBI, Maslach y Jackson, 1981)

Teacher Attitude Scale (TAS, Farber, 1984)

Teacher Burnout Scale (Seidman y Zager, 1986)

Burnout Measure (Pines y Aronson, 1988)

Efectos Psíquicos del Burnout (García-Izquierdo, 1990)

Matthews Burnout Scale for Employees (Matthews, 1990)

Escala de Variables Predictoras del Burnout (Aveni y Albani, 1992)

Cuestionario de Burnout del Profesorado (Moreno, Oliver y Aragonese, 1993)

Tedium Occupational (TO, Corcoran, 1995)

General Survey (MBI-GS, Schaufeli, Leiter, Maslach y Jackson, 1996)

Cuestionario Breve del Burnout (CBB, Moreno, Bustos, Matallana y Miralles, 1996)

Role Hassles Index (Zohar, 1997)

Cuestionario de Desgaste Profesional de Enfermería (CDPE, Moreno, Garrosa y González, 2000)

Cuestionario de Burnout Profesorado (CBP-R, Moreno, Garrosa, y González, 2001)

Oldenburg Burnout Inventory (OLBI, Demerouti, Bakker, Vardakou y Kantas, 2003)

Cuestionario para la evaluación del Síndrome de Quemarse en el Trabajo (CESQT, Gil-Monte, 2005)

Cuestionario Urgente de Burnout (CUBO, De la Gandara, García, González-Corrales y Álvarez, 2005)

Copenhagen Burnout Inventory (CBI, Kristensen, Borritz, Villadsen Christensen, 2005)

Cuestionario Desgaste Profesional Médico (CDPM, Moreno, Gálvez, Garrosa y Mingote, 2006)

c. MOBBING

Leymann Inventory of Psychological Terrorization (LIPT, Leymann, 1990a)

Work Harassment Scale (WHS) (Björkqvist et al., 1994)

Bergen Bullying Index (Einarsen, Raknes y Matthiesen, 1994)

Negative Acts Questionnaire (NAQ, Einarsen y Raknes, 1997)

Workplace Incivility Scale (WIS, Cortina, Magley, Williams y Langhout, 2001)

Escala LIPT reducida (Van Dick y Wagner (2001)

Cuestionario para el Estudio del Acoso Laboral en Contextos Universitarios (CEALCU, Justicia et al., 2002)

MOBB-90; MobbCF-21; MobbCG-15 y MobbCS-28 (escalas elaboradas por Boada, De Diego y Vigil, 2003)

LIPT adaptación española (González de Rivera y Rodríguez-Abuín, 2003)

NAQ-RE, revisión española de la versión reducida (García-Izquierdo et al., 2004)

Barómetro Cisneros (Fidalgo y Piñuel, 2004)

Generalized Workplace Harassment Questionnaire (GWHQ, Rospenda y Richman, 2004)

Uncivil Workplace Behaviour Questionnaire (UWBQ, Martin y Hine, 2005)

CAPP Cuestionario de Acoso Psicológico Percibido (Morán, 2007)

Cuestionario de Acoso Psicológico en el Trabajo (CAPT, Moreno et al., 2008)

Cuestionario de Hostigamiento en el Trabajo (HPT, Fornés, Martínez-Abascal y García de la Banda, 2008)

NAQ-R (Einarsen, Hoel y Notelaers, 2009)

Los cuestionarios utilizados para la evaluación de riesgos psicosociales pretenden investigar la relación individuo-organización del trabajo, tal y como es percibida y experimentada por los trabajadores, cuya finalidad es la de identificar condiciones de trabajo que llevan implícita la probabilidad de causar un daño (riesgo), medir sus probabilidades y adoptar las medidas preventivas necesarias. Esto es así porque la evaluación psicosocial efectuada al amparo de la LPRL pretende identificar y valorar riesgos psicosociales. En ningún caso tiene como objeto evaluar el estrés, la salud, el bienestar, la satisfacción o motivación laboral de los trabajadores; en todo caso, estas medidas sí serían de gran utilidad utilizadas de manera complementaria (Niño, 2006). Por tanto, es un error utilizar otro tipo de escalas para la medida de riesgos psicosociales, tal y como asevera la ITSS (2012):

No se puede optar por algún cuestionario o escala cuya finalidad no es la detección de riesgos, por ejemplo escalas de clima o satisfacción. Éstas pueden ser utilizadas como información complementaria o para correlacionar con los datos sobre riesgos guía actuación inspectora (p. 25).

Con todo, es habitual encontrarse con cuestionarios de evaluación psicosocial que sobrepasan su cometido y van más allá de la evaluación de las condiciones de trabajo, adentrándose en el estado de salud y bienestar de los trabajadores e incluso evaluando la *salud* de la propia organización (ej., clima laboral) (véase tablas 13 a 20). Este tipo de actuaciones no hace sino acrecentar la confusión que rodea la evaluación psicosocial, dando lugar en más de una ocasión a situaciones de intromisión profesional (Niño, 2002, 2006). Cabe recordar que la identificación de riesgos con objeto de eliminarlos o minimizarlos es la esencia de la denominada *prevención primaria*, competencia de los técnicos de prevención de riesgos laborales. Sin embargo, los chequeos sobre la salud, entre los que se encuentran las evaluaciones de estrés, así como su diagnóstico y tratamiento, corresponden a la *vigilancia de la salud*, parcela exclusiva del personal sanitario cualificado (art. 22.6 LPRL y 37.3 RSP).

Proposición 7:

Los cuestionarios de evaluación de riesgos psicosociales no deben incluir escalas cuya finalidad no sea detectar riesgos, como escalas de satisfacción o clima laboral

Moreno y Báez (2010) recomiendan seguir una serie de criterios a la hora de valorar la conveniencia de utilizar uno u otro instrumento de medición, (1) grado de difusión en publicaciones de calidad sobre el comportamiento psicométrico de la herramienta y sobre la eficiencia del proceso de evaluación tras la intervención; (2) baremación en una muestra amplia con población nacional; (3) información psicométrica, incluida la información sobre la validación factorial confirmatoria; (4) accesibilidad del instrumento, teniendo en cuenta que el acceso gratuito propicia el fomento del proceso de evaluación y que los manuales y la documentación de apoyo facilitan el uso autónomo de los cuestionarios; y (5) comparación con otros métodos de evaluación, en aras de conocer sus ventajas y limitaciones.

De todos los instrumentos recogidos en la Tabla 12, se han seleccionado para un estudio detallado los elaborados en España e incluidos en la segunda categoría -los que informan acerca de un conjunto de factores psicosociales-, ya que son las herramientas más utilizadas en el

ámbito de prevención de riesgos laborales (CEOE, 2013; ITSS, 2012; Moreno y Báez, 2010; Peiró et al., 2013), pues (1) son aptas para realizar una evaluación inicial o global; y (2) no deberían sobrepasar el ámbito profesional de los técnicos de prevención. Por otra parte, las técnicas específicas como las escalas de *burnout* o *mobbing* son apropiadas como medidas reactivas (ITSS, 2012).

Se han establecido como criterios de inclusión en la revisión incluida en la presente tesis las técnicas de evaluación de riesgos contempladas en la *Guía de actuaciones de la Inspección de trabajo y seguridad social sobre riesgos psicosociales* (ITSS, 2012), que a su vez cumplan con los requisitos establecidos por Guàrdia (2010a, 2010b) y Meliá (2008), es decir, sólo se tendrán en cuenta los instrumentos de los que exista información social y acerquen un mínimo de información técnica y teórica (ver Tablas 13 a 20)¹.

Tabla 13: Revisión FPSICO 3.1 (2014)
Fuente: Elaboración propia a partir de Moreno y Báez (2010)
FPSICO 3.1 (Martín-Daza y Pérez, 1997) INSHT

FPSICO 3.1 (Martín-Daza y Pérez, 1997) INSHT				
Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Uso libre: www.insht.es	89-Likert	9	Teoría clásica de los Test	Teorías de estrés, motivación y satisfacción
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
- Tiempo de trabajo	1660	.697	<u>Validez de criterio:</u> “r” entre cada factor y variables como: (1) Escala General de Satisfacción (Warr, Cook y Wall en 1979) (2) Escala de Salud General (GHQ, Goldberg (1972) Según los autores, se obtuvieron mayoritariamente, correlaciones moderadas y elevadas.	a) Cuestionario previo a la evaluación para recogida de información.
- Autonomía	1455	.865		b) Cuestionario de 89
- Carga de trabajo	1593	.733		ítems tipo Likert,
- Demandas psicológicas	1465	.737		agrupados en 9 factores.
- Contenido/variedad	1539	.705		c) Aplicación informática
- Desempeño de rol	1549	.732		de resultados y
- Supervisión/participación	1556	.844		recomendaciones
- Interés por el	1582	.842		
trabajador/Compensación	1520	.716		
- Relaciones y apoyo social				
Validación según Ferrer, Guílera y Peiró (2011)				

* α de Cronbach

r: Correlación de Pearson

El FPSICO es de muy reciente actualización (2014), sin embargo, el estudio de las propiedades psicométricas corresponde a la versión anterior (Ferrer et al., 2011) donde el número de ítems era menor. Se han encontrado estudios donde la consistencia interna arroja valores pobres (ej., Callejón-Ferre, Montoya-García., Pérez-Alonso y Rojas-Sola, 2015; Meseguer et al., 2007). Por otra parte, Peiró et al. (2013) afirman que la evidencia empírica publicada en la literatura científica es limitada. El cuestionario es de fácil uso y está baremado en una muestra amplia de población utilizando percentiles, además cuenta con estudios

¹ En la primera versión de este apartado también se examinó la herramienta del Instituto Navarro de Salud Laboral (INSL) para la identificación y evaluación de los factores psicosociales (Lahera y Góngora (2002; Lahera y Nogareda, 2009). No obstante, dicha herramienta ha sido descatalogada recientemente (Instituto de Salud Pública y Laboral de Navarra, ISPLN, 2016), optando entonces por eliminarla de este estudio.

comparativos con otros métodos, sobre todo con el CoPsoQ-ISTAS 21. Como inconveniente, Moreno y Báez (2010) indican un pobre grado de difusión en publicaciones científicas rigurosas.

Tabla 14: Revisión Cuestionario Multidimensional DECORE
Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).

Cuestionario Multidimensional DECORE (Luceño, 2005; Luceño, Martín, Tobal et al., 2005) Universidad Complutense de Madrid.				
Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Uso libre previo pago: http://web.teaediciones.com	44-Likert	4	Teoría clásica de los Test	- Modelo Demanda-Control-Apoyo social de Karasek (1979) y Johnson y Hall (1988) - Modelo de Esfuerzo-Recompensa de Siegrist (1996a)
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
- Demandas cognitivas	299	$\alpha \geq .79$	<u>Validez de criterio:</u> ANOVA de un factor	a) Cuestionario de 44 ítems tipo Likert, agrupados en 4 factores.
- Recompensas	315	$\alpha \geq .82$	Variables criterio: estrés, satisfacción, baja médica por enfermedad, fatiga y accidentes laborales.	b) Aplicación informática de corrección online a través de la web de TEA ediciones.
- Apoyo organizacional			41% de la varianza explicado a través de los 4 factores.	
- Autonomía			<u>Validez de constructo:</u> Análisis factorial exploratorio y confirmatorio muestra buen ajuste	
Validación según Luceño, Martín, Díaz et al., 2008) y Luceño et al. (2010)				

* α de Cronbach
ANOVA: Análisis de varianza

El DECORE proviene del ámbito académico y es comercializado a través de TEA Ediciones. Es accesible previo pago sólo para psicólogos, y los técnicos de prevención interesados tendrán que acreditar conocimientos técnicos en la materia y titulación. Por lo tanto, no es una herramienta de libre acceso. Sus autores recomiendan el uso de esta herramienta para muestras masivas. Según la revisión efectuada por la CEOE (2011), se centra demasiado en los trabajadores y poco en la organización evaluada, además de no proporcionar orientaciones sobre las medidas preventivas. Fue diseñado para utilizar en el ámbito clínico, y en consecuencia sobrepasa las competencias de la ergonomía y psicología aplicada desde el punto de vista de la prevención de riesgos psicosociales. Aunque es fácil de utilizar, no utiliza la terminología con propiedad. Según los criterios establecidos por Moreno y Báez (2010), la herramienta no contempla una baremación efectuada sobre una muestra amplia, y tampoco contribuye al fomento de la evaluación dado su carácter comercial; sin embargo, su grado de difusión en publicaciones científicas es satisfactorio.

Tabla 15: Revisión CoPsoQ-ISTAS 21

Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).

CoPsoQ-ISTAS 21 (versión 2, Moncada, Utzet et al., 2014)				
Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Uso libre Www. istas.net/copsoq/m	30-Likert (v. corta) 109-Likert (v. media)	15 20	Teoría clásica de los Test	- Modelo Demanda-Control-Apoyo social de Karasek (1979) y Johnson y Hall (1988) - Modelo de Esfuerzo-Recompensa de Siegrist (1996a) - Teoría del capital social (Cook y Wall, 1980; Elovainio, Kivimäki, Steen y Vahtera, 2004)
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
- Doble presencia	5100	.89	<u>Validez de criterio:</u>	a) Cuestionario en papel para cada versión.
- Exigencias cuantitativas		.69	La mayoría de escalas	b) Tablas resumen.
- Ritmo de trabajo		.76	correlacionan en la	c) Circulares informativas sobre el proceso en todas sus etapas.
- Exigencias emocionales		.76	dirección esperada con las	d) Matrices de exposición, origen y medidas preventivas.
- Exigencia de esconder emociones		.62	5 dimensiones de salud	e) La versión media contiene programa informático con manual de uso.
- Influencia		.77	(satisfacción en el	
- Posibilidades de desarrollo		.81	trabajo, salud general,	
- Sentido del trabajo		.74	salud mental, estrés,	
- Apoyo social de los compañeros		.83	burnout) excepto	
- Apoyo social de superiores		.87	Influencia. En Moncada,	
- Calidad de liderazgo		.90	Utzet et al., 2014 se	
- Sentimiento de grupo		.89	explican los motivos de	
- Previsibilidad		.59	mantener dicha escala.	
- Claridad de rol		.78	No se incluyen datos	
- Conflicto de rol		.75	específicos sobre la	
- Reconocimiento		.79	validez de constructo,	
- Inseguridad sobre el empleo		.62	aunque los autores dicen	
- Inseguridad sobre condiciones de trabajo		.78	que el análisis factorial	
- Justicia		.81	cuenta con buen ajuste	
- Confianza vertical		.72		
Validación según Moncada, Utzet et al., 2014				

* α de Cronbach.

El CoPsoQ-ISTAS 21 es una de las pocas herramientas validadas y estandarizadas en una amplia muestra de trabajadores (n=7612 en el año 2005 y n=5100 en el año 2010). Constituye la mayor muestra española hasta la fecha y es la única que muestra valores de referencia para cada una de las dimensiones (Moncada, Llorens, Font, Galtés y Navarro, 2008). Otra característica diferenciadora del resto es que se actualizan periódicamente los valores poblacionales de referencia, al menos una vez cada cinco años (Moncada, Utzet et al. 2014). Cuenta con alto grado de difusión en publicaciones científicas nacionales e internacionales (Moncada, Utzet et al., 2014). Es de uso libre y gratuito y se han realizado estudios comparativos con otras herramientas de evaluación de riesgos psicosociales (ej., FPSICO). Por lo tanto, cumple con todos los criterios propuestos por Moreno y Báez (2010) a excepción de ofrecer datos sobre su estructura factorial.

La nueva versión ha eliminado, modificado y añadido escalas; concretamente se agrega la dimensión de capital social (Kristensen, 2010). Se utiliza el modelo epidemiológico, donde se

definen hasta siete unidades de análisis, como puesto de trabajo, sexo o tipo de contrato, y permite compararlas con su referente poblacional. Se puede adaptar el cuestionario a la realidad de cada organización.

Tabla 16: Revisión INERMAP

Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).

INERMAP: Educación (EDUMAP); Industria (PSICOMAP); Sanidad (SANIMAP); Oficinas administración (OFIMAP); Bancos (OFIMAP BANCA); Comercios (OFIMAP COMERCIALES); Teleoperadores (TELEMAP); PYMES (PSICOPYME) - Métodos del Instituto de Ergonomía MAPFRE (INERMAP) (Gracia-Camón, 2003)				
Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Uso libre previo pago www.inermap.com/software/ Psicofaq.htm#P2 Principales	Variable según escala	Variable dependiendo del sector a evaluar	Instrumento desarrollado por validez de contenido directo, no se encuentran datos sobre el modelo psicométrico utilizado.	Según el autor se basa en varias teorías y modelos explicativos de cada uno de los riesgos psicosociales a medir.
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
Medidas comunes a todas las escalas: - Distribución y diseño ambiental - Comunicación - Control y Liderazgo - Carga mental - Turnicidad - Satisfacción social	?	$\alpha > .92$	validez de contenido y de constructo $> .85$	Técnicas cualitativas y cuantitativas a) Observación directa. b) Cuestionario c) Entrevista d) Check-list
Los escasos de datos de validación han sido publicados en la web de la organización y se refieren al INERMAP en general: http://www.inermap.com/ . No se facilitan datos sobre fuentes ni sobre el tamaño muestral utilizado para obtener el índice de fiabilidad, ni se aporta información sobre los procedimientos seguidos en la validez de constructo ni de contenido				

* α de Cronbach

En cuanto a los cinco criterios reseñados por Moreno y Báez (2010), este método no cumple ninguno, ya que sólo está disponible mediante pago. Su difusión suele hacerse mediante publicaciones en instituciones propias (MAPFRE). No se encuentran publicaciones ni otro tipo de fuentes donde comprobar las garantías psicométricas (más allá de las cifras que el autor asegura que corresponden a la fiabilidad y a la validez). Tampoco está baremado y no se han encontrado estudios comparativos con otros métodos.

Así, los responsables del método justifican la ausencia de validación del método alegando que el cuestionario no es la evaluación de riesgos (Gracia, *s.f.*) sino que sólo aporta información adicional y por lo tanto no es necesario realizar pruebas de fiabilidad y validez. Su método de evaluación da mayor importancia a la observación del técnico, relegando a un segundo plano la evaluación subjetiva mediante cuestionario. De hecho, la respuesta de Gracia (*s.f.*) cuando los usuarios del método solicitaron información psicométrica de la medida fue decir que: “si existe validez NECESARIAMENTE ha de existir fiabilidad” (p. 2) y por lo tanto solicitar ambos parámetros es redundante. No es de extrañar que para algunos investigadores la herramienta no contempla las condiciones mínimas de rigor psicométrico (Guardià, 2008; Moreno y Báez, 2010).

Tabla 17: Revisión MC Mutual-UB

Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).

Batería de Evaluación de Riesgos Psicosociales para las PYMES de MC Mutual y Universitat de Barcelona (Guàrdia et al., 2008). Universidad de Barcelona y MC Mutual

Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Sólo para empresas asociadas a la mutua http://www.mc-mutual.com	46-Likert	7	Test referido al criterio	No se basan en modelos teóricos sino en criterios de aplicabilidad de prevención de riesgos. Así, las dimensiones que evalúan fueron decididas por sesiones entre 6 jueces expertos en la materia.
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
- Organización del tiempo de trabajo	821	.747	<u>Validez de constructo:</u> AFC con buen ajuste, a excepción del índice de bondad c^2 AFE muestra buen ajuste	Técnicas cualitativas y cuantitativas
- Comunicación		.849		a) Cuestionario de 46 ítems tipo Likert, agrupados en 7 factores.
- Formación y desarrollo		.817		b) Entrevista
- Efectos sociales y de grupo		.881		c) <i>Check-list</i>
- Participación		.895		b) Aplicación informática
- Contenido del trabajo		.844		
- Exigencias de la tarea y del entorno de trabajo		.569		
		.951 (Global)		
Validación según Guàrdia et al. (2008)				

* α de Cronbach

AFC: Análisis Factorial Confirmatorio

AFE: Análisis Factorial Exploratorio

MC Mutual-UB es una técnica de fácil aplicación que combina metodología cualitativa y cuantitativa para favorecer la triangulación, que es la característica más destacable del método. Está pensada para PYMES y microempresas y contiene medidas preventivas generales que tendrán que ser adaptadas a la realidad de cada organización. Es también una de las pocas herramientas que contiene medidas de control interno de sesgos, concretamente deseabilidad social y sinceridad (Guàrdia, 2008).

Volviendo a hacer referencia a las recomendaciones para seleccionar las técnicas de evaluación idóneas (Moreno y Báez, 2010), el MC Mutual-UB sólo cumple el requisito de proporcionar información psicométrica muy completa, incluyendo las cargas factoriales para cada dimensión.

Tabla 18: Revisión Prevenlab Psicosocial
Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).

METODO PREVENLAB PSICOSOCIAL (MODELO AMIGO) UNIVERSIDAD DE VALENCIA (Peiró, 1999b, 2006)				
Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Uso por expertos, no es libre ni gratuito	118-Likert	21 (agrupados en 11 dimensiones)	Teoría clásica de los Test	<p>Modelo teórico propio: MODELO AMIGO</p> <p>Se basa sobre todo en modelos de estrés teniendo como estresores:</p> <ul style="list-style-type: none"> - Desequilibrio entre demandas-control - Desajuste entre demandas y recursos. - Discrepancias entre realidad y valores personales - Intensidad del estímulo - Apreciación primaria y secundaria-respuesta emocional y afectiva - Estrategias de afrontamiento - Efectos - Consecuencias
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
<ul style="list-style-type: none"> - Sistema de trabajo y actividad y condiciones de trabajo en su puesto - La tecnología que utiliza en su trabajo - La estructura de la empresa - Infraestructura y recursos disponibles en la empresa - Clima social y comunicación en la empresa - Políticas y prácticas de RRHH en la empresa - Estilos de dirección y relaciones con directores y supervisores - Personas y grupos de trabajo en la empresa - Misión y cultura de la empresa - La estrategia de la empresa - El ambiente en el que actúa la empresa 	550	.87 .80 .84 .73 .93 .91 .95 .89 .81 .84 .90	<p><u>Validez de criterio</u></p> <p>Correlaciones en el sentido esperado</p> <p><u>Validez de constructo:</u></p> <p>AFE avala la estructura multifactorial</p>	<p>Técnicas cualitativas y cuantitativas</p> <p>a) Cuestionario de 118 ítems tipo Likert, agrupados en 11 factores.</p> <p>b) Entrevista</p> <p>c) Sesiones de grupo</p> <p>Analiza los riesgos implementando 3 módulos:</p> <p>1) módulo de chequeo general de la empresa</p> <p>2) módulo de profundización en el diagnóstico de los riesgos</p> <p>3) módulo <i>Surveyfeedback</i> (devolución de la información)</p>
<p>Validación según Peiró (2006)</p> <p>Se facilitan los resultados de la validez factorial, sin embargo no hacen eco de los procedimientos utilizados. Tampoco se encuentran datos sobre su baremación.</p>				

* α de Cronbach

AFE: Análisis Factorial Exploratorio

El método Prevenlab, Modelo AMIGO (Modelo de Análisis Multifacético para la Intervención y Gestión Organizacional) es idóneo para la evaluación inicial general y posteriores evaluaciones específicas (Peiró, 2006). En cuanto a su forma de acceso, no es una herramienta gratuita. Además, es restringida pues sólo puede ser aplicada por especialistas en el método que cumplan determinados requisitos. Como indica el propio autor, estamos ante una

metodología que tiene que ser aplicada por personal con “conocimientos y formados en la práctica de la consultoría experta y la consultoría para el cambio organizacional” (Peiró, 2006, p. 115). Esta condición puede escapar del alcance de una gran parte de técnicos de prevención de riesgos psicosociales. Al hilo de las limitaciones con las que se pueden encontrar los mencionados técnicos de prevención, el método AMIGO mide dimensiones que abarcan aspectos que exceden de las condiciones laborales; además, utiliza terminología propia de la psicología de la Gestalt y de otras disciplinas psicológicas, incurriendo, de ese modo, en falta de propiedad a la hora de utilizar el lenguaje de los conceptos preventivos (CEOE, 2013).

En cuanto a las características técnicas, no se demuestra correlación significativa entre varios factores y el promedio del total de los riesgos. Entre los factores con baja significación se encuentran (1) las condiciones ambientales y demandas físicas del trabajo, (2) aspectos relacionados con el uso o estado de la tecnología del trabajo, (3) comunicación en la empresa, (4) distintos estilos de dirección, (5) perspectivas estratégicas que implican cambio y (6) posibles cambios radicales.

En cuanto a los puntos a tener en cuenta a la hora de seleccionar una herramienta eficaz de evaluación, según Moreno y Báez (2010) estamos ante una metodología con un alto grado de difusión en publicaciones rigurosas, aportando información sobre las propiedades psicométricas, si bien no cumple con los requisitos de baremación ni con el de fomentar y facilitar las evaluaciones de riesgo psicosocial al ser un instrumento de acceso restringido.

Tabla 19: Revisión MARC-UV

Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).

MÉTODO MARC-UV				
A: BATERIA DE FACTORES PSICOSOCIALES DE SALUD LABORAL				
B: BATERIA VALENCIANA PREVACC				
A) Batería de factores psicosociales de salud laboral (BFPSL) Universidad Valencia (Meliá, 2003a)				
Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Uso por expertos, no es libre ni gratuito http://www.uv.es/seguridadlaboral	154-Likert	13	Teoría clásica de los Test	No se especifica
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
- Conflicto de rol	513	.91	Validez de criterio	Técnicas cualitativas y cuantitativas
- Ambigüedad de rol		.91	“r” moderadas y elevadas	a) cuestionario de 154 ítems tipo Likert, agrupados en 13 factores.
- Exposición externa		.91	entre:	b) otras técnicas cualitativas opcionales
- Mobbing		.86	(1) variables de consecuencias psicológicas	
- Liderazgo		.83	(2) siniestralidad laboral con los factores psicosociales de la batería	
- Comunicación organizacional		.86		
- Calidad, Seguridad y medio ambiente		.83		
- Satisfacción laboral		.94		
- Salud psicológica		.85		
- Tensión asociada al trabajo		.93		
- Burnout		.81		
- Propensión al abandono		.82		
- Rendimiento		.78		
Validación según Meliá (2004a).				
No se hallan datos sobre la validez factorial				

Continuación tabla 19

B) Batería Valenciana PREVACC (Meliá, 2003b)				
Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
No libre	77-Likert	6	Teoría clásica de los Test	- Modelo psicosocial de los accidentes laborales Meliá (1998, 2006) - Investigaciones sobre la importancia y contribución del factor humano Heinrich (1931), Seo (2005)
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
- Exposición a riesgos - Probabilidad percibida de accidentes - Respuesta de seguridad de la empresa - Respuesta de seguridad del trabajador - Respuesta de seguridad compañeros - Respuesta seguridad supervisores	513	$\alpha > .79$	Validez de criterio: "r" con buen ajuste entre: (1) las escalas entre si (2) indicadores de accidentes de trabajo (3) Batería BFPSL	Técnicas cualitativas y cuantitativas a) entrevista a directivos b) cuestionario de 77 ítems tipo Likert, agrupados en 6 factores para los trabajadores c) cuestionario para el técnico o investigador d) cuestionario de información objetiva y subjetiva para el técnico o investigador
Validación según Meliá (2004b) No se hallan datos sobre la validez factorial				
* α de Cronbach r: Correlación de Pearson				

Este método consta de dos baterías que pueden utilizarse complementariamente o por separado: (a) Batería de Factores Psicosociales de Salud Laboral y (b) Batería Prevacc. La primera de ellas contiene factores psicosociales tradicionales y factores que no pueden considerarse estrictamente condiciones de trabajo (ej., *mobbing*, propensión al abandono, *burnout*), siendo fácil de aplicar y de rápida administración, e incluye además una medida para el control de la deseabilidad social. Guàrdia (2010b) la describe como herramienta de diseño obsoleto, centrada en los trabajadores en detrimento de los factores clave de la organización evaluada. No es una herramienta gratuita y presenta versiones traducidas a varios idiomas, lo que ha contribuido a su difusión internacional.

Respecto a la segunda batería, PREVACC, no se trata de una herramienta de evaluación de riesgos psicosociales al uso. Según Meliá (2004b) "es un conjunto articulado de instrumentos sobre los aspectos de factor humano directamente ligados con la seguridad" (p. 12). Moreno y Báez (2009, 2010) la describen como metodología compleja. Al igual que en la batería BFPSL no se encontraron datos sobre la validez factorial y, según CEOE (2013), utiliza términos complejos difíciles de apreciar desde la esfera preventiva.

Tabla 20: Revisión WONT

Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).

Metodología WONT- Cuestionario RED (Recursos, Experiencias y Demandas)

Equipo de investigación WONT (*Work and Organizacional Network*) de la Universidad Jaume I de Castellón. (Salanova et al., 2006)

Versiones:

RED TIC para usuarios de tecnologías de la información y de la comunicación/ RED TT para teletrabajadores
RED UNI para estudiantes universitarios / RED ADI para evaluar adición al trabajo.

Accesibilidad	Nº Ítems	Nº Factores	Modelo psicométrico	Fundamentación teórica
Acceso restringido al equipo WONT www.wont.uji.es	150-Likert	37	Teoría clásica de los Test	Modelo Demanda- Control- Apoyo social de Karasek (1979) y Johnson y Hall (1988) - Modelo de Esfuerzo- Recompensa de Siegrist (1996a) - Modelo de Demandas- Recursos (Demerouti, Bakker, Nachreiner, y Schaufeli, 2001) - Modelo de Proceso Dual (Salanova, Bresó, y Schaufeli, 2005; Schaufeli y Bakker, 2004).
Medidas principales	n	Fiabilidad*	Validez	Herramientas del método
- Demandas - Recursos laborales - Recursos personales - Daños psicosociales - Bienestar psicosocial - Daños organizacionales	?	de .70 a .96	Análisis factorial con buen ajuste Los autores remiten a numerosas publicaciones lo que dificulta hacerse una idea de las características globales de validez de sus escalas.	Técnicas cualitativas y cuantitativas a) <i>Check-list</i> b) entrevista c) cuestionario de 150 ítems tipo Likert, agrupados en 37 factores d) Grupos de discusión e) Aplicaciones on-line
Validación según Salanova et al. (2006)				

* α de Cronbach

Esta metodología implica un proceso largo y complejo, encuadrado dentro de la psicología ocupacional positiva, y por lo tanto el foco de atención recae sobre la promoción de la salud en detrimento de la prevención de riesgos. Desde esta perspectiva, la promoción de la salud, aun siendo una *cualidad favorable*, como los autores la denominan, se distancia de los planteamientos legales de la prevención de riesgos psicosociales y por ello alejados también de las competencias de los técnicos de prevención de riesgos laborales (CEOE, 2013; Moreno y Báez, 2010). Prueba de ello es que no sólo se evalúan los riesgos sino también las consecuencias psicológicas y los daños para la organización (ej., aburrimiento, ansiedad, *burnout*, relajación, absentismo o intención de abandono entre otras variables). Al igual que con el método DECORE, la CEOE (2013) le reprocha la falta de propiedad en el uso de la terminología, conceptos y normativa preventiva. Su uso es limitado al propio equipo WONT y se ha difundido tanto en publicaciones nacionales como internacionales.

En las publicaciones que tratan el método WONT no se encuentran descripciones completas de las variables que mide, de tal forma que, las revisiones efectuadas por Guàrdia,

(2010b) y Moreno y Báez (2010) alegan falta de claridad en los datos que se refieren a la validación, y echan de menos descripciones detalladas de los datos psicométricos, pues se trata de una metodología compleja. Respecto a la estructura factorial confirmatoria existen pocos estudios.

Una vez revisados los diferentes mecanismos disponibles para la evaluación psicosocial lo apropiado sería disponer de un balance crítico sobre la pertinencia de utilización de unos u otros instrumentos. Haciendo eco de los argumentos y análisis propuestos por Moreno y Báez (2010), se presenta a continuación una tabla comparativa (tabla 21) con los principales criterios que, según estos autores, deberían de tenerse en cuenta a la hora de seleccionar una herramienta de evaluación psicosocial.

Tabla 21: Criterios de selección de la herramienta adecuada
Fuente: Adaptado de Moreno y Báez (2010)

	Grado de difusión en publicaciones científicas de reconocido prestigio	Muestra amplia de baremación	Información sobre estructura factorial	Fomento de la evaluación mediante uso libre y gratuito	Estudios comparativos entre instrumentos
FPSICO V3.1		✓*	✓	✓	✓
CoPsoQ -ISTAS 21 V2	✓	✓		✓	✓
DECORE	✓		✓		
PREVENLAB	✓				
INERMAP					
MARC-UV	✓				
WONT RED	✓				
MC MUTUAL			✓		

* La baremación corresponde a la versión anterior (V.3.0) realizada por Ferrer et al. (2011).

Como se ha mencionado en reiteradas ocasiones a lo largo de esta revisión teórica, los cuestionarios de autoinforme son las herramientas que se utilizan con más frecuencia para evaluar los riesgos psicosociales en el ámbito laboral, existiendo a tal efecto test específicos y con propiedades psicométricas satisfactorias, que han sido validados y adaptados con muestras españolas (Meliá, 2008). Su fiabilidad y validez, sumada a la ausencia de sesgo, es la información que debe incorporar cualquier test, siguiendo las directrices recomendadas desde el ámbito psicométrico (Abad et al., 2011). A pesar de ello, y según el análisis expuesto en la tabla 21, solamente dos herramientas cumplen con la mayoría de criterios, el F-PSICO y CoPsoQ (Cuestionario Psicosocial de Copenhague). Ambas han sido baremadas con una muestra amplia de población, han sido comparadas con otros instrumentos de evaluación, fomentan la evaluación al ser de acceso libre y gratuito, y cuentan con un manual de uso autónomo. A diferencia del CoPsoQ, el F-PSICO dispone de información sobre su estructura factorial; sin embargo, el grado de difusión en literatura científica es mínimo si se compara con la proyección internacional del CoPsoQ-ISTAS 21.

En cuanto a la ausencia de sesgos, y tal y como se desprende de la revisión efectuada, son pocos los cuestionarios que contemplan o hacen referencia explícita al control de sesgos; como excepción destacan la batería MC-Mutual y la batería Prevacc. En este sentido, llama la

atención que en las exhaustivas revisiones efectuadas por diversos investigadores (CEOE, 2013; Guàrdia, 2008b; Moreno y Báez, 2010) se pone mayor énfasis en la necesidad de que los autores de las escalas den a conocer la fiabilidad y estructura factorial de las herramientas que de explicitar si contienen o no alguna medida de control de sesgos.

Según estos planteamientos, para poder comprobar la calidad psicométrica de los cuestionarios de evaluación, los técnicos de prevención de riesgos psicosociales deberían poseer ciertos conocimientos acerca del grado de fiabilidad, validez, estandarización, objetividad e influencia de sesgos de las escalas, lo cual les facilitaría la elección de la mejor herramienta de evaluación para la organización o situación que se pretenda evaluar. Sin estos conocimientos, comparar el rigor científico entre herramientas o corregir los posibles sesgos de repuesta, se antoja difícil.

Proposición 8:

Actualmente apenas existen herramientas para evaluar riesgos psicosociales que contengan control de sesgos.

Proposición 9:

A la hora de aplicar un test suele darse más importancia a las propiedades psicométricas de fiabilidad y validez que a la detección de sesgos.

Proposición 10

Es necesario tener conocimientos de psicometría para interpretar los resultados de las escalas de autoinforme utilizadas para valorar los riesgos psicosociales.

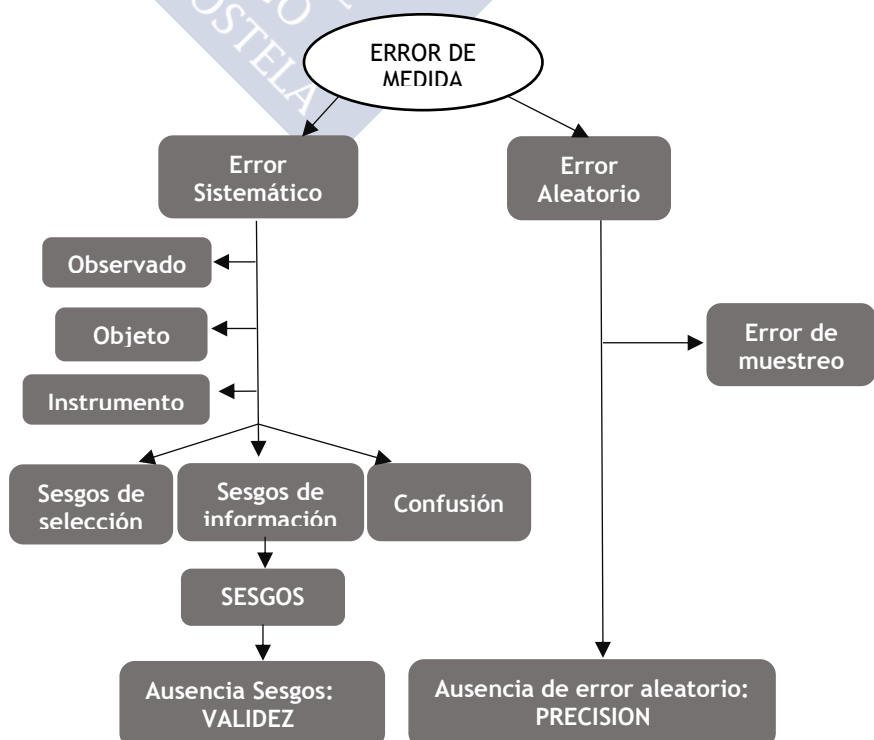
3. APLICACIÓN, CORRECCION E INTERPRETACION DE RESULTADOS DE LAS MEDIDAS DE AUTOINFORME

3.1. AMENAZAS A LA VALIDEZ EN LAS MEDIDAS DE AUTOINFORME

Cuando se aplica un método de evaluación de riesgos psicosociales, como el recomendado en la NTP 450, se está siguiendo el método científico; esto es, un protocolo de acciones sucesivas y técnicas que nos llevarán a obtener un conocimiento. Una de las principales fases de este protocolo corresponde a la selección y aplicación de los instrumentos de investigación adecuados, y bajo ese marco, las herramientas más utilizadas son las escalas de autoinforme (ver epígrafe 2.4). Este tipo de escalas, están a menudo compuestas por ítems de respuesta cerrada, lo que podría propiciar escalas sesgadas debido a la influencia que ejercen este tipo de ítems en el modo de responder de los encuestados (Ferrando, Lorenzo-Seva, 2010; Jensen, 1995; Weijters, Geuens y Schillewaert, 2010; Tomás, Sancho, Oliver, Galiana y Meléndez, 2012).

Las fuentes de error pueden tener su origen en el objeto o sujeto a medir, en el mismo instrumento de medida utilizado y/o en la persona que realiza la medición (De la Guardia, Gómez, Sandoval y García, 2012; Restrepo y Gómez- Restrepo, 2004). Tradicionalmente, estas amenazas se catalogan en error aleatorio y error sistemático (ver figura 5). Aunque ambos conceptos fueron introducidos en epígrafes anteriores (véase el epígrafe 2.3.4, *Rigor científico de los test*), en las líneas que siguen se tratan con mayor profundidad.

Figura 5: Errores de medida
Fuente: Elaboración propia

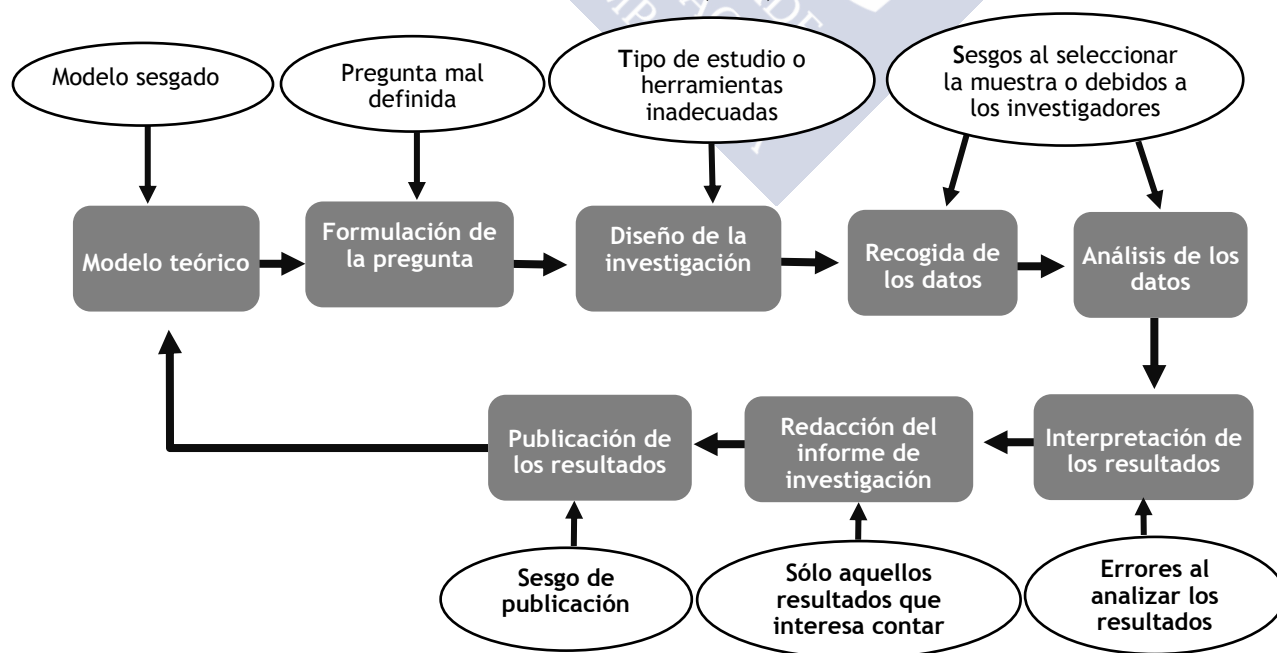


Se entiende por error aleatorio o accidental, aquella parte de nuestra experiencia que no se puede predecir (Rothman, 2012). Son incertidumbres debidas a factores casuales que afectan a la medición, pero no pueden atribuirse al contenido del ítem o a factores del informante. Por ejemplo, instrucciones ambiguas, distracciones momentáneas o equivocaciones del encuestado a la hora de marcar la respuesta correcta. Estos errores son imposibles de evitar ya que toda medición contiene cierto grado de incertidumbre (Abad et al., 2011; Mendoza, 2012; Nunnally y Bernstein, 1995). Este tipo de error es de carácter variable, no sigue ningún patrón definido y cada vez que se repita la medida en las mismas condiciones el error variará aleatoriamente, unas veces sobrestimando y otras subestimando (Navas, 2010). De este modo, cuanto menor sea el error aleatorio mayor será la precisión de los resultados. Por ese motivo está directamente relacionado con la fiabilidad de la herramienta de medida. Se obtendrá un indicador fiable cuando, aplicado repetidas veces el mismo instrumento en las mismas condiciones, los resultados arrojen puntuaciones idénticas o muy similares (Nunnally y Bernstein, 1995).

Aunque no es posible eliminar la varianza aleatoria, si se pueden realizar esfuerzos para minimizarla. Por una parte, se pueden intentar mantener uniformes las condiciones de la prueba, controlando por ejemplo las instrucciones, el entorno o el tiempo, entre otros elementos (Nunnally y Bernstein, 1995). Por otra parte, se puede aumentar el tamaño de la muestra ya que a mayor tamaño mayor precisión, de ahí la importancia del tamaño muestral en las investigaciones (Broeck, y Brestoff, 2013; De la Guardia et al., 2012; M. Hernández, Garrido, y Salazar, 2002; Porta, 2014; Salamanca, 2013). Dado que el error aleatorio es inherente a toda investigación, se hace preciso estimar su magnitud ya que pueden restar peso a las generalizaciones de los resultados (Navas, 2010).

El otro tipo de error es el error sistemático, también denominado sesgo, causado por algún factor que se ha obviado (véase figura 6).

Figura 6: Diferentes errores y sesgos en las diferentes etapas del proceso investigador
Fuente: Salamanca (2013)



Gómez-Benito et al. (2010) delimitan el concepto como sigue:

El sesgo se refiere a la injusticia derivada de uno o varios ítems del test al comparar distintos grupos que se produce como consecuencia de la existencia de alguna característica del ítem o del contexto de aplicación del test que es irrelevante para el atributo medido por el ítem (p. 76).

A grandes rasgos, los sesgos se deben a fallos en el diseño, conducción o análisis de un estudio, dando lugar a desviaciones constantes en los resultados cada vez que se realiza, ya sea sobre los mismos participantes o sobre distintos sujetos (M. Clark. y Oxman, 2003). En contraposición al error aleatorio, el sistemático sigue siempre una dirección determinada: o presenta valores más altos o más bajos que los reales (Corbetta, 2003; Cortada de Kohan, 1999). Si se conoce su causa, en buena parte podría evitarse, corregirse o compensarse. El sesgo afecta a la validez y no se reduce al aumentar el tamaño muestral, pues es un error producido en el diseño del estudio, al seleccionar la muestra o al medir las variables. A todos los efectos, su magnitud tiene que determinarse de forma empírica, considerando adecuados los cuestionarios que tengan menos del 15% de error sistemático (Reidl, Guillén, Sierra y Joya, 2002). Se clasifica en tres tipos, sesgos de selección, de información y de confusión (Broeck, y Brestoff, 2013; De la Guardia et al., 2012; M. Hernández et al., 2002; Porta, 2014; Salamanca, 2013).

El sesgo de selección o muestreo se produce al elegir una muestra que no es representativa de la población que se pretende estudiar; o que contiene variables homogéneas cuando la población sobre la que queremos generalizar los resultados es heterogénea (Corbetta, 2003; Salamanca, 2013). El sesgo de información, por su parte, es resultado de imprecisiones en la medición de las características de los sujetos y, como consecuencia, los participantes del estudio están incorrectamente clasificados, afectando tanto a la validez interna como externa (Abad et al., 2011; Schoenbach y Rosamond, 2000). Por último, el sesgo de confusión o *confusor*, se diferencia de los otros dos, por el hecho de que los datos obtenidos pueden ser verdaderos, no obstante, son atribuidos erróneamente a la variable de estudio cuando en realidad son debidos a la influencia de una o más variables externas denominadas variables de confusión (Schoenbach y Rosamond, 2000). De esta forma, la distorsión puede producirse subestimando, sobreestimando e incluso revirtiendo la dirección del resultado de interés (Aguayo, 2007). En este sentido, uno de los confusores hallados en correlaciones estrés-salud ha sido el *afecto negativo*, cuya influencia espuria actúa inflando los resultados (Watson y Pennebaker, 1989).

Si bien, toda investigación es susceptible de contener sesgos, dependiendo de su diseño, serán más frecuentes un tipo de sesgos que otros. Sirva de referencia Sackett (1979), quien elaboró una lista de los sesgos más frecuentes en estudios analíticos. En la misma línea se encuentra el catálogo de sesgos de Delgado-Rodríguez y Llorca (2004), donde se aporta un completo inventario de sesgos ordenados alfabéticamente, especificando su tipología (selección, información, confusión) e indicando los tipos de estudios afectados por los mismos.

Tal y como se presentó en la figura 6, los errores pueden aparecer en cualquier fase de la investigación (Moya, 2009; Salamanca, 2013). Incluso antes de su comienzo es primordial tener claro qué se va a evaluar y para qué. La validez de la investigación comienza a fraguarse mucho antes de construir las herramientas de medida, pues unos fundamentos teóricos erróneos dejarán sin sustento la validez de constructo (AERA, et al., 2014; Muñiz, 2004; Padilla, Gómez,

Hidalgo y Muñiz, 2006; Smith, 2005). El tipo de herramientas a utilizar es otro punto crítico, así como una clara delimitación del contexto, circunstancias y población sobre la que se va a medir, ¿se van a utilizar encuestas o cuestionarios?, ¿de lápiz y papel u *online*?, ¿será individual o colectivo?, ¿los resultados servirán de diagnóstico, de orientación o de selección? (AERA et al., 2014; Zumbo, 2007). No tener despejadas estas cuestiones es dejar abierta la puerta a posibles sesgos. En cuanto a la construcción del instrumento de medida, la redacción de los ítems es de suma importancia, tiene que ser especialmente cuidadosa, pues de ella depende, en gran medida la calidad del cuestionario y, por ende, de toda la investigación. Algunos aspectos a tener en cuenta en la formulación de los reactivos es la representatividad, la extensión, la claridad de los enunciados o también la deseabilidad social (AERA et al., 2014; R.J. Cohen y Swerdlik, 2006; Morales, 2006, 2011, 2012; Muñiz, Fidalgo, García-Cueto, Martínez y Moreno, 2005; D. T. Morse y Morse, 2002; Prat y Doval, 2005). Todos los esfuerzos hasta aquí realizados para evitar los sesgos pueden resultar infecundos si la muestra seleccionada no es la adecuada en tamaño, heterogeneidad o representatividad (Broeck, y Brestoff, 2013; Corbetta, 2003; De la Guardia et al., 2012; M. Hernández et al., 2002; Porta, 2014).

Aun siguiendo todos los principios teóricos y métricos, debe confrontarse la variabilidad humana. Es decir, (1) los encuestados que pueden malinterpretar las instrucciones de la prueba, equivocarse al marcar las respuestas, responder de modo socialmente deseable, pueden estar influenciados por las condiciones ambientales (ruido, calor, frío) del lugar donde realizan la prueba (Muñiz y Bartram, 2007; Muñiz et al., 2005; Nunnally y Bernstein, 1995) o pueden incluso mentir (Martínez-Selva, 2005); (2) los investigadores pueden equivocarse al realizar los análisis y/o malinterpretar los resultados e incluso subestimar los resultados no deseados (R. Hernández et al., 2014; Morris y Maisto, 2005); y (3) revisores y directores de revistas científicas que en ocasiones aceptan la difusión de investigaciones pueden dejarse llevar por la significación o por la direccionalidad de los resultados en detrimento de los estudios menos significativos. De hecho, se estima que un 50% de las investigaciones de diferentes disciplinas no se publica nunca (Bosques y Gómez, 2009; Higgins y Green, 2011).

Desde AERA, et al. (2014) y la Comisión Internacional de test (ITC, 2010) han recomendado cumplir una serie estándares de buenas prácticas relacionadas con el uso de los test en investigación. Entre uno de sus objetivos se encuentra guiar a los investigadores hacia estudios libres de sesgos y, para ese fin, proponen seguir varias directrices que pasan por el entrenamiento previo del observador o investigador, así como de preparar con antelación el lugar de la prueba cerciorándose de ser cómodo y estar libre de distracciones; procurar establecer una relación de empatía con los encuestados e intentar reducir sus niveles de ansiedad, o por no menos no reforzarla; eliminar los posibles focos de desviación de la atención como alarmas, teléfonos móviles o leer las instrucciones pausada y claramente, entre otras muchas recomendaciones.

Proposición 11:

Un protocolo de administración donde, por ejemplo, los encargados de administrar los test lean las instrucciones pausada y claramente y traten de reducir la ansiedad de los evaluados, contribuiría a minimizar los posibles sesgos de respuesta.

Tal y como se ha visto, los sesgos pueden darse en cualquier momento de la investigación, pero los sesgos que interesan para esta tesis, son los sesgos propios de los test de autoinforme. En este sentido, son numerosas las variables que pueden propiciar que las puntuaciones del test no sean justas, destacando los aspectos socioculturales, de género, de construcción y adaptación del test, el estilo y tendencia de respuesta (Abad et al., 2011; Cheung y Rensvold, 1999; Delgado, 1998; Espelt, Doval, y Viladrich, 2014; Espelt, Viladrich et al., 2014; Restrepo y Gómez-Restrepo, 2004). En esa línea, se considera que las medidas de autoinforme son sensibles a numerosos factores de sesgo como la aquiescencia, la deseabilidad social, los errores escalares (estilo de respuesta central y severidad), el olvido, las distorsiones de memoria, la ignorancia, el cansancio, la falta de motivación, y un largo etcétera (Barker et al., 2002; Cronbach, 1946, 1950; De las Cuevas y González de Rivera, 1992; Delgado, 1998; Donaldson y Grant-Vallone, 2002; Echeburúa et al., 2003; González e Iruarrizaga, González- Llaneza, 2005, 2007; Jensen, 1995; F. Jiménez y Sánchez, 2004; Matesanz, 1997; Morales, 2006; Nunnally y Bernstein, 1995; Paulhus, 1991; Restrepo y Gómez-Restrepo, 2004; Sackett, 1979; Tourangeau y Rasinski, 1988).

Respecto a sesgos propios de autoinformes utilizados en la evaluación de riesgos psicosociales, no se ha encontrado información específica en la literatura sobre prevención de riesgos laborales, teniendo que utilizar como referente estudios sobre escalas utilizadas con otros objetivos, tales como la medición de variables cognitivas, de personalidad y de salud. En definitiva, todas las escalas de autoinforme, incluidas las de evaluación de riesgos psicosociales, son medidas subjetivas que tienen en común la medición de la realidad percibida por el sujeto y es esta percepción la que puede estar modulada por diversos elementos como la deseabilidad social, el autoengaño o la simulación (Rodríguez-Abuín y González de Rivera, 2006). A modo de muestra, se recogen en la tabla 22 los sesgos que algunos autores atribuyen a las escalas de autoinforme utilizadas en psicología.

Tabla 22: Sesgos y Fuentes de error en escalas de autoinforme

Fuente: Elaboración propia

EVALUACIÓN PSICOLÓGICA (De las Cuevas y González de Rivera, 1992)

Efecto halo
Indulgencia e inclemencia
Tendencia central o extrema
Restricción de rango y efectos de frecuencia
Cambio de respuesta
Falsificación
Deseabilidad social
Sesgo izquierda-derecha
El error de contraste
La aprehensión evaluada

QUE AFECTAN A GRUPOS MINORITARIOS (Delgado, 1998)

Sesgo de género
Contenido inapropiado
Lenguaje inapropiado
Formato inapropiado
Muestras de estandarización inapropiadas
Condiciones de aplicación inapropiadas

Continuación tabla 22

PSICOLOGIA CLINICA FORENSE (Echeburúa, Amor y Corral, 2003)	
TENDENCIA DE RESPUESTA	
Aquiescencia	
Tendencia de respuesta central o extrema	
Ignorancia	
Falta de motivación	
Simulación o disimulación voluntaria	
Deseabilidad social	
Distorciones de memoria	
EVALUACION PSICOLOGICA (González e Iruarrizaga, 2005)	
DISTORSIONES DE RESPUESTA	
Estilos de repuesta sobredimensionados	Patrones de mala imagen
Simulación	
Estilos de respuesta infra dimensionados	Patrones de buena imagen
Defensividad	
Deseabilidad social	
Disimulación	
Aquiescencia	Estilo de respuesta aleatorio
No aquiescencia	
EVALUACION PSICOLOGICA (González-Llaneza, 2007)	
Los efectos de la situación	
Efectos del examinador	
Simulación	
Deseabilidad social	
Tendencias de respuesta	
Estilos de respuesta	
EVALUACION CONDUCTUAL (Jensen, 1995)	
Deseabilidad social	
Características de la población	
Características de exigencia de la evaluación	
Influencias reactivas	
Sesgos de respuesta	
Las instrucciones	
Formato de respuesta	
Redacción de ítems	
Variables relativas al lugar donde se cubre el cuestionario	
Número de instrumentos de medida administrados simultáneamente	
Exposición a experiencias de sensibilización	
Interacciones con el evaluador	
EVALUACION PERSONALIDAD (F. Jiménez y Sánchez, 2004)	
Defensividad	
Fingimiento positivo	
Fingimiento negativo	
Responder al azar	
Deseabilidad social	

Otro ejemplo ilustrativo se encuentra es la revisión realizada por Choi et al. (2010) que, a modo de catálogo, recoge 48 fuentes de error propias de los cuestionarios de salud. Gran parte de su interés radica en que se ha centrado únicamente en los sesgos específicos de los cuestionarios, que pueden ser comunes a todo tipo de escalas de autoinforme. Puesto que su cotejo podría resultar de gran utilidad para la elaboración y/o evaluación de cuestionarios, se incluye en la tabla 23 la denominación de los sesgos revisados por dichos autores.

Tabla 23: Catálogo de sesgos clasificados por fuentes

Fuente: Adaptado de Choi et al. (2010)

CATÁLOGO DE SESGOS CLASIFICADOS POR FUENTE DE LOS MISMOS EN CUESTIONARIOS DEL ÁREA DE SALUD		
A. SESGOS DERIVADOS DE PROBLEMAS EN LA REDACCIÓN DE LA PREGUNTA	A.1. Redacción no clara de la pregunta	A.1.1 Redacción ambigua A.1.2 Pregunta muy compleja o larga A.1.3 Dos o más pregunta en un solo enunciado
	A.2 Palabras no bien definidas o de uso poco común	A.2.1 Palabras poco comunes o jerga técnica A.2.2 Palabras cuyo significado no es preciso
	A.3 Disociación entre pregunta y objetivo	A.3.1 Creencias vs conocimientos A.3.2 Período de tiempo inadecuado para objetivo
	A.4 No se colectan datos necesarios	A.4.1 Datos secundarios
	A.5 Problemas con escalas o categorías	A.5.1 Escala con poca capacidad discriminatoria
		A.5.2 Selección forzada (categorías insuficientes)
		A.5.3 Incongruencia entre preguntas y respuestas
		A.5.4 Intervalo faltante
		A.5.5 Intervalos superpuestos
		A.5.6 Tendencia central
	A.6 Preguntas guiadoras	A.6.1 Estructural
		A.6.2 Focalización
		A.6.3 Cadencia
B. SESGOS DERIVADOS DE PROBLEMAS CON EL DISEÑO Y DIAGRAMACIÓN DEL CUESTIONARIO	A.7 Preguntas invasoras	A.7.1 Autodefensa
		A.7.2 Tema delicado
	A.8 Preguntas muy cortas	A.8 Preguntas muy cortas
	A.9 Preguntas sin consistencia	A.9.1 Definición de caso
		A.9.2 Cambio de escala
		A.9.3 Cambio en las palabras usadas en la redacción
		A.9.4 Diagnostico vago
	A.10 Negociación en enunciado	A.10.1 Negociación en enunciado
	B.1 Diagramación inadecuada	B.1.1 Diagramación de respuestas
		B.1.2 Escalas yuxtapuestas
		B.1.3 Alineación respuestas
	B.2 Cuestionarios demasiado largos	B.2.1 Decidir entre No o Si
		B.2.2 Fatiga
	B.3 Interpretación de respuestas abiertas	B.3.1 Respuesta abierta
	B.4 Estructura del cuestionario	B.4.1 Saltos en el cuestionario

Continuación tabla 23

C. SESGOS DERIVADOS DEL PROBLEMA CON EL USO DEL CUESTIONARIO	C.1 Falta de objetividad del entrevistador	C.1.1 Entrevistador C.1.2 Expectativa
	C.2 Reacción psicológica del respondedor	C.2.1 Aversión a los extremos C.2.2 Satisfacción positiva
	C.3 Reacciones conscientes psicológicas	C.3.1 Respuestas falsas y premeditadas C.3.2 Aceptabilidad social C.3.3 Enfermedad estigmática C.3.4 Rechazo social o cultural C.3.5 Sospecha de causa
	C.4 Aprendizaje de respondedores	C.4.1 Aprendizaje C.4.2 Anticipación de la hipótesis
	C.5 Recuerdo defectuoso	C.5.1 Mínimo esfuerzo C.5.2 Fuente secundaria C.5.3 Memoria C.5.4 Telescópico
	C.6 Barrera sociocultural	C.6.1 Sociocultural

3.2. SESGOS DE RESPUESTA

De entre la diversidad de distorsiones que afectan a los cuestionarios de autoinforme, los sesgos de respuesta son posiblemente, los de mayor incidencia. Llevan casi un siglo siendo objeto de estudio en las ciencias sociales. Se considera que el inicio de las investigaciones se debe a Allport y Hartmann (1925, citado por Cantril, 1946) y desde entonces se han sucedido innumerables publicaciones (ej., Adams, 1961; Alwin y Krosnick, 1991; Bachman y O'Malley, 1984; Barnette, 2000; Batchelor, Miao y McDaniel, 2013; Chami-Castaldi, Reynolds y Wallace, 2008; Cronbach, 1946, 1950; Dupuis et al., 2015; Edwards, 1953; Hamilton, 1968; Hui y Triandis, 1985, 1989; Jackson y Messick, 1958; Jensen, 1995; Likert, 1932; R.E. McGrath, Mitchell, Kim y Hough, 2010; Messick, 1966, 1968; Nardi, 2003; Nunnally y Bernstein, 1995; Paulhus, 1981, 1984, 1986, 1991; Paulhus y Vazire, 2007; Rorer, 1965; Swearingen, 1997, 1998; Viswanathan, 2005).

Este tipo de sesgos pueden ser deliberados o inconscientes, llegando a ser un verdadero problema para determinadas investigaciones, pues contaminan los resultados de los instrumentos de medida, pudiendo invalidar parte de los resultados o incluso toda la evaluación (Baumgartner y Steenkamp, 2001; Delgado, 1998; Dupuis et al., 2015; Embree y Whitehead, 1991; Harzing, Brown, Köster y Zhao, 2012; Hui y Triandis, 1989; Marquis, Marquis y Polich, 1986; Stening y Everett, 1984; Tomás et al., 2012).

Proposición 12:

Los sesgos son un problema importante pues podrían contaminar los resultados de medida de los test, pudiendo incluso invalidar toda la evaluación

Para Paulhus (1991) “un sesgo de respuesta es una tendencia sistemática a responder a una parte de los ítems de una escala en base a alguna razón distinta al contenido específico del ítem” (p. 17). Cuando se responde a un ítem, lo preferible es pensar que la respuesta depende exclusivamente de su contenido, sin embargo, entran en juego otros factores como su formulación o la personalidad del que responde (Morales, 2006). Por ejemplo, “un individuo puede estar más inclinado a responder *sí* o *verdadero* en lugar de *no* o *falso* en una prueba de respuestas breves” (R. J. Cohen y Swerdlik, 2006, p. 346), también puede contestar inducido por la presión del tiempo para responder el cuestionario, o en caso de duda entre verdadero-falso o de acuerdo-en desacuerdo, puede existir una tendencia a estar siempre de acuerdo. También puede responder marcando la opción más deseable socialmente de entre otras alternativas (Messick, 1991).

Cronbach (1946, 1950) fue el primero en dar una importancia significativa a estas formas de responder. Encontró que las medidas de personalidad, actitud, interés y capacidad son más sensibles a estos modos de respuesta, ocasionando que personas con el mismo nivel de atributo recibiesen diferentes puntuaciones, lo que sin duda afecta a la validez de la medida. Para referirse a este tipo de respuestas utilizó el término *response set* (tendencia de respuesta), definiéndolo como “el hábito o predisposición temporal que impulsa a una persona a responder a los ítems del test de modo distinto a como lo haría si el mismo contenido fuese presentado de diferente forma” (1946, p. 476). No obstante, observó que algunas tendencias de respuesta se mantenían de manera habitual a lo largo del tiempo, pudiendo ser un reflejo de aspectos de la propia personalidad del encuestado (Cronbach, 1946; Messick, 1968). En consecuencia, Jackson y Messick (1958) propusieron utilizar el término *response style* (estilo de respuesta) para referirse a los sesgos de respuesta que se muestran estables en el tiempo, para diferenciarlos de los *response set*, que son temporales. Rorer (1965) utilizó definiciones diferentes para referirse a los mismos conceptos: “El término *set* connota un consciente o inconsciente deseo por parte de un entrevistado para responder de tal forma que se presente una cierta imagen de sí mismo” (p. 133), mientras que “*style* se refiere a seleccionar alguna categoría de respuesta durante una cantidad de tiempo desproporcionada independientemente del contenido del ítem” (p. 134).

Son muchos los autores que coinciden en agrupar los sesgos de respuesta bajo las denominaciones de tendencia respuesta y estilo respuesta, entre ellos Harzing et al. (2012), van Herk, Poortinga y Verhallen (2004), Kulas y Stachowski, (2009), Man (2013), Messick, 1991; Paulhus (1984, 1986, 1991) o Peer y Gamliel (2011). Si bien, la definición mayormente citada es la ofrecida por Paulhus (1991, p. 17):

Un sesgo de respuesta puede ser una tendencia de respuesta, que es, una reacción temporal a una demanda situacional, por ejemplo, la presión del tiempo o las expectativas sociales. Alternativamente, un sesgo puede ser inducido por los efectos del contexto como el formato del ítem o la naturaleza de ítems anteriores en el cuestionario (para una revisión, ver Tourangeau y Rasinski, 1998). En la medida que un individuo manifieste el sesgo consistentemente a través del tiempo y las situaciones, se dice que el sesgo es *estilo de respuesta* (Jackson y Messick, 1958; Wiggins, 1973).

Más recientemente, Peer y Gamliel (2011) señalan que:

Estilo de respuesta es la tendencia a distorsionar las respuestas en una dirección particular más o menos independientemente del contenido del estímulo. *Tendencia de respuesta* es el deseo consciente o inconsciente por parte del encuestado para contestar de manera que se produzca una cierta imagen de sí mismo (p. 1).

En conclusión, los sesgos de respuesta pueden ser debidos: (i) al cuestionario en sí, por ejemplo, ítems redactados de forma ambigua; (ii) a una predisposición estable para responder a los ítems de una escala de una determinada manera independientemente del contenido (estilo respuesta); y (iii) a una forma de responder motivada por una reacción temporal, incluyendo las respuestas que tratan de ofrecer, según las situaciones, una imagen positiva o negativa del encuestado, ocultando o distorsionando la información (tendencia de respuesta).

La tendencia de respuesta más popular y estudiada es la deseabilidad social. Cano-García, García-Martínez, Rodríguez-Franco y Antuña-Bellerín (2005) se refieren a ella diciendo que “consiste en la tendencia a elegir comportamientos socialmente aceptables: indican buena integración social, pero puntuaciones extremas se interpretan como falta de sinceridad...” (p. 71). Aunque también hay otros estilos de respuesta habituales como la aquiescencia, la tendencia central de respuesta y la severidad o tendencia extrema de respuesta (Harzing et al. 2012; Kulas y Stachowski, 2009; Navas, 2010, véase Man, 2013 para una revisión más exhaustiva de los diferentes estilos de respuesta). Cano-García et al. (2005), los definen como sigue:

La aquiescencia se refiere a la tendencia a responder en el sentido que plantea el ítem, con independencia de su contenido; se asocia con actitudes conformistas o negativistas ante la prueba. La tendencia central se refiere a la ignorancia de las calificaciones de los extremos de los ítems, mientras que en la tendencia extrema se ignoran las puntuaciones centrales de la escala; ambas pueden indicar errores en la comprensión, escasez de significación de los ítems o rigidez personal (pp. 70-71).

De la revisión efectuada se puede concluir que las distorsiones más influyentes en los autoinformes son los sesgos de respuesta referidos a la aquiescencia y la deseabilidad social (Anastasi y Urbina, 1998; Barker et al., 2002; Cheung y Chan, 2002; Cronbach, 1976; Ferrando y Anguiano-Carrasco, 2012; Harzing et al., 2012; Nawas, 1971; Paulhus, 1991; Polaino-Lorente y Martínez Cano, 2003; Stricker, 2013), pudiendo actuar, incluso, de forma simultánea. Lo que propicia que en la presente tesis se les dedique mayor atención.

Proposición 13:

Los sesgos de respuesta son posiblemente los de mayor incidencia en los cuestionarios de autoinforme

3.2.1. Tendencia de respuesta: Deseabilidad Social

La deseabilidad social (DS) es la variable que ha recibido más atención (Cosentino y Castro, 2008) y es definida como la tendencia a responder de acuerdo con estándares sociales aceptables en la comunidad (Collazo, 2005; Crowne y Marlowe, 1960; Edwards, 1953, 1957, 1964; Ellingson, Smith y Sackett, 2001; González e Iruarrizaga, 2005; Nederhof, 1985;

Paulhus, 1984). En ocasiones se utilizan otros apelativos dentro de la etiqueta de DS, como *sinceridad*, *distorsión de respuesta*, *simulación*, *virtudes infrecuentes* o *distorsión motivacional* (Salgado, 2005). Contribuye a distorsionar de manera sutil la imagen que la persona da de sí misma al responder los ítems del cuestionario de la manera socialmente más deseable (Li y Bagger, 2007; Nuevo et al., 2009; Nunnally y Bernstein, 1995). Para Martínez Selva (2005) la distorsión de la información, la mentira, el engaño y la ocultación son conductas inherentes al ser humano, y entre las numerosas motivaciones que expone, se encuentra el propio beneficio o beneficiar a otros, ofrecer una imagen positiva ante la sociedad, evitar castigos o el bien social. Por lo tanto la distorsión puede ser positiva o negativa (Goffin y Boyd, 2009; Mueller-Hanson, Heggstad y Thornton, 2003). La DS es uno de los sesgos que más preocupación e interés ha suscitado, sobre todo en la evaluación psicológica medida con autoinformes, debido a ello son muy prolíficos los trabajos dedicados a su estudio, pues como afirma De las Cuevas y González de Rivera (1992):

Los inventarios auto-administrados están especialmente sujetos a la simulación o falsificación por parte de quien los cumplimenta. La mayoría de los ítems de dichos inventarios tienen una respuesta reconocible como socialmente más aceptable o deseable que las demás (...). Parece, por tanto, fácil crear deliberadamente la impresión que se desea en estos inventarios (p. 3).

Proposición 14

Las medidas de autoinforme son susceptibles de ser falseadas

El estudio de este sesgo comenzó a finales de los años 20 con Hartshorne y May (1928), para poco después retomar las investigaciones Bernreuter (1933) y Vernon (1934). Desde entonces, el estudio de la DS ha experimentado un crecimiento sin parangón, que se ha ido enmarcando entre dos vertientes diferenciadas:

- a) La DS es considerada como una tendencia de respuesta que incrementa el error de medida y por lo tanto afecta a la validez de las medidas de autoinforme (Cronbach, 1946; Crowne y Marlowe, 1964; Edwards, 1953, 1957, 1990; H. J. Eysenck y Eysenck, 1976; Furnham, 1986; Hartshorne y May, 1928; Hathaway y McKinley, 1943; Jackson, 1986; Meehl y Hathaway, 1946; Paulhus, 1991). Esta es la postura más extendida.
- b) La DS es un rasgo estable de personalidad vinculado a la necesidad de aprobación y a la conformidad social y no como un sesgo (Adams, Ryan, Hoffman, Dobson y Nielsen, 1985; Allaman, Joyce y Crandall, 1972; Dicken, 1963; Evans, 1982; Ferrando y Chico, 2000; Fleming y Zizzo, 2011; Hogan y Nicholson, 1988; Jacobson, Kellogg, Cauce y Slavinet, 1977; McCrae y Costa, 1983; Millham y Kellogg, 1980; Paulhus y John, 1998; Pauls y Stemmler, 2003; Sackeim y Gur, 1978; Salgado, Remeseiro e Iglesias, 1996).

Entre ambas posturas, destaca el cuidadoso estudio de Borkenau y Ostendorf (1989) quienes exploraron la relación entre distintas variables de personalidad y la tendencia de responder con DS, hallando correlación inversa entre el neuroticismo y la tendencia a dar respuestas socialmente deseables. Hay que aclarar que esta nueva consideración como rasgo de

personalidad no implicó que el sesgo dejase de evaluarse, sino que se modificó su criterio de interpretación (Poo, 2007). No obstante, la polémica volvió a retomarse cuando se encontraron correlaciones bajas entre distintas herramientas de medida de DS (Stöber; Dette y Musch, 2002), y cuando Paulhus (1984, 2002) propuso la bidimensionalidad del constructo de DS, distinguiendo dos dimensiones: *Impression Management* (manejo de impresiones) y *Self-deception* (autoengaño).

Uno de los objetivos de la evaluación psicológica es poder tener una medida libre de distorsión y, a pesar de existir posturas discordantes, la DS (sea sesgo o sea rasgo, de una dimensión o dos) es una fuente de invalidez cuando directamente “se requiere que las personas digan algo acerca de ellas mismas” (Enríquez y Domínguez, 2010, p. 70). En la literatura ha quedado patente efecto de la DS en los cuestionarios de personalidad, con importantes repercusiones en la práctica clínica, forense y médico-legal (González e Iruarrizaga, 2005), sin dejar de lado el campo de la Psicología del Trabajo y de las Organizaciones, donde también pesan sus efectos (Salgado, 2005).

Relacionado con el estrés laboral, Semmer, Grebner y Elfering (2004) corroboraron la susceptibilidad de los autoinformes a posibles asociaciones espurias entre las condiciones del trabajo y el bienestar laboral. Éstas estaban originadas precisamente por los sesgos de respuesta como la deseabilidad social y la afectividad negativa. Estos postulados son reforzados por Gamero-Burón y González (2013) cuando interpretan los altos índices de estrés laboral recopilados de varias encuestas (Encuesta de calidad de vida en el trabajo, ECVT; Encuesta nacional de salud, ENSE, 2006 y del *International Social Survey Programme*, ISSP, 1997), donde más del 60% de la población asalariada a nivel nacional reportaba sufrir estrés. Según estos autores “podría estar presente el sesgo de deseabilidad social en las respuestas: los trabajadores podrían estar inclinados a indicar que experimentan estrés en la medida en que esto implica que el trabajador es un miembro valioso de la organización” (Gamero-Burón y González, 2013, p. 97). Otra anotación importante a la que hacen alusión es la posibilidad de que en contexto de crisis económica los trabajadores declaren mayores niveles de estrés para así dotar de más valor su trabajo, pues muchos otros carecen de él. En la misma dirección se expresa Duro (2005). Apunta que la tendencia al sesgo será positiva o negativa dependiendo de la naturaleza del contenido del ítem, como ejemplo expone que “variables relativas a síntomas psicológicos o consecuencias laborales negativas -v.g., absentismo- se obtendrán distribuciones asimétricas positivas; y para variables relativas a rasgos, consecuencias o actuaciones del trabajador más positivas -v.g., capacidad personal general- tenderán a resultar distribuciones asimétricas negativas” (p. 61).

Con todo, el número de trabajos empíricos que aborda específicamente el estudio de la DS en prevención es reducido. Sirva de ejemplo el estudio de Louzán (2014), en el que se exploró el impacto de los factores psicosociales en una muestra de 201 trabajadores del sector de la construcción de la provincia de Lugo, utilizando el método CoPsoQ-ISTAS 21. En concreto, se observó que los niveles de exposición más desfavorables para la salud (45%) se hallaron en el factor de doble presencia. Curiosamente, el 97% de los participantes fueron hombres, siendo la doble presencia “específico de las mujeres” (Moncada et al., 2007; M. Rivas, 2008, p. 231). El resultado tendría su explicación en la distorsión de las respuestas con objeto de mostrar una imagen socialmente aceptable. Afirmación que concuerda con los trabajos de Achen y Stafford

(2005), Kan (2008), Press y Townsley (1998) y Sigelman, Budd, Spanhel y Schoenrock (1981), quienes encontraron discrepancias al alza en las puntuaciones de los hombres, cuando se les preguntó sobre su participación en las tareas domésticas.

Otra explicación posible sobre la distorsión consciente de las respuestas en los cuestionarios de evaluación de riesgos puede estar relacionada con el contrato psicológico. Este alude “al conjunto de creencias y expectativas acerca de los compromisos y obligaciones percibidas en una relación, más allá de los acuerdos formales manifiestos” (Vesga, 2007, p. 81). Este acuerdo de naturaleza informal tiene componentes tanto implícitos como explícitos (Schein, 1982) y desde el punto de vista del trabajador, abarca entre otros “aspectos como ascensos, trato interpersonal, estabilidad laboral, formación, apoyo en momentos de crisis, etc.” (Vesga, 2011, p. 174). La influencia del contrato psicológico se ha relacionado directamente con la satisfacción laboral, el compromiso con la organización, la evaluación positiva de las relaciones laborales, la permanencia en la empresa (Guest, 2004, Guest y Conway, 2002), la confianza, la lealtad y el rendimiento (Dabos y Rousseau, 2004; Saldarriaga, 2013; Solanes y Zaragoza, 2010; Tena, 2002). Gracia et al. (2006) lo consideran explicativo de actitudes y comportamientos y una de las posibles conductas del empleado, cuando percibe que se ha vulnerado o violado dicho contrato, es dañar a la empresa para contrarrestar el perjuicio padecido por las expectativas incumplidas (Bies y Tripp, 2005; Robinson y Rousseau, 1994). No hay que olvidar que la evaluación de riesgos laborales es una imposición legal para la empresa y una evaluación negativa no favorece en nada a la organización, entre otros motivos por el posible impacto económico al tener que adoptar medidas para corregir las incidencias detectadas. Entonces podría no ser descabellado pensar que trabajadores descontentos con la organización puedan falsear sus respuestas con objeto de perjudicar a la empresa, y por el mismo razonamiento, trabajadores afines a la empresa puedan dar respuestas favorecedoras.

Proposición 15:

La percepción de incumplimiento de contrato psicológico, por parte del trabajador, podría influir en el falseamiento consciente de las respuestas de escalas utilizadas en la evaluación de riesgos psicosociales.

3.2.2. Estilos de respuesta: Respuesta Aquiescente

La aquiescencia es otro de los sesgos de respuesta más estudiados y que ha generado gran discusión. En términos generales, es entendida como la preferencia a contestar afirmativamente sin tener en cuenta el contenido (Billiet y Davidov, 2008; Couch y Keniston, 1960; Cronbach, 1946; Echeburúa et al., 2003; Lentz, 1938; Martin, 1964; Matesanz, 1997; Tomás et al., 2012; Viswanathan, 2005). Según Nawas (1971), “significa, en esencia, que una persona responde un cuestionario o inventario como ‘verdadero’ o ‘falso’ o ‘de acuerdo’ o ‘en desacuerdo’ posiblemente sin pensar mucho en la sustancia del ítem, particularmente cuando el sujeto está acosado por la duda” (p. 202). Couch y Keniston (1960), administrando cientos de ítems de una larga lista de escalas, demostraron dos tipos de tendencias de respuesta, los que siempre están de acuerdo con los ítems *Yeasayers* y los que siempre están en desacuerdo *Naysayers*. En la misma línea, otros autores también realizaron la misma diferenciación, agrupando en un lado los modos de responder con predisposición a responder *si*, *verdadero* o *de acuerdo*, en

contraposición a las tendencias de respuesta negativas, ya sea *no*, *falso* o *en desacuerdo* (Anastasi, 1988; Baumgartner y Steenkamp, 2001; Paulhus, 1991).

A principios del siglo XX, otros investigadores como Cogan, Conklin y Hollingworth (1915, citado por Ferrando, Condon y Chico, 2004), Fritz (1927) y Mathews, (1929, citado por Morales, 2006) emprendieron los primeros estudios, pero no fue hasta Cronbach (1946, 1950) cuando comenzaron a adquirir significado. Sin embargo, con la revisión crítica de Rorer (1965) la aquiescencia fue considerada un mito, teniendo que esperar hasta las investigaciones de Ray en la década de los 70, para alcanzar un gran auge, reavivando de nuevo la polémica. A juicio de Morales (2006) fue Ray quien le ha dado la importancia debida a la aquiescencia.

La gran controversia creada acerca de la aquiescencia no pone en duda su existencia, pues multitud de evidencias han puesto de manifiesto que “hay sujetos que tienden a responder afirmativamente en caso de duda, y que en parte se contradicen en sus respuestas (eso es precisamente la aquiescencia)” (Morales, 2006, p. 157). Si no que la polémica gira alrededor de su valoración, ¿qué peso tiene la aquiescencia en la interpretación de los resultados de las mediciones? Este es un debate todavía inacabado, algunos autores la consideran insignificante (ej., Gove y Geerken, 1977; Grimm y Church, 1999, Kerlinger, 1975; Nunnally y Bernstein, 1995; Rorer, 1965; Rorer y Goldberg, 1965; Wright, 1975) y otros afirman que la aquiescencia puede ser fuente de contaminación en el resultado de las mediciones (Jackson y Messik, 1961; Ray, 1983). Así, Ray (1990) cita más de 50 referencias donde se considera la aquiescencia como un problema a tener en cuenta en las medidas de personalidad y actitudes.

Otro enfoque presente en el problema de la respuesta aquiescente trata sobre si catalogarla como un rasgo de personalidad (Couch y Keniston, 1960; Gough y Heilbrun, 1983; Messick, 1966; Nunnally y Bernstein, 1995). Según Morales (2006) no existen datos consistentes:

Se han verificado múltiples correlaciones con variables de personalidad que en su conjunto no ofrecen pruebas consistentes de que efectivamente, se trate de un rasgo; más bien parece que se trata de un tipo de respuesta ante determinados estímulos concretos. Se podría categorizar la aquiescencia como rasgo de personalidad si las correlaciones entre diversas medidas de aquiescencia fueran altas y si las medidas de aquiescencia pudieran predecir *conductas aquiescentes*. De nada de esto hay datos medianamente concluyentes (p. 158).

Según varios estudios, la tendencia a la aquiescencia es más patente en escalas unidireccionales, en ítems demasiado largos y complejos y en escalas de actitud en comparación con las de personalidad (Billiet y McClendon, 1998; Ferrando, Demestre, Anguiano-Carrasco y Chico, 2011; Ferrando y Lorenzo-Seva, 2010; Finlay y Lyons, 2002; van Herk et al., 2004; Krenz y Sax, 1987; Morales, 2006; Paulhus, 1991; Schuman y Presser, 1981). Se ha comprobado que también estimulan la respuesta aquiescente (1) los ítems formulados en sentido positivo en comparación a los formulados en sentido inverso (Elliot, 1961; Morales, 2006; Schriesheim y Hill, 1981), (2) el grado de ambigüedad de los ítems (Banta, 1961; Bass, 1955; Cronbach, 1950; Finlay y Lyons, 2002; Lemos, 2005; Messick, 1967; Morales, 2006; Ray, 1983; Viswanathan, 2005); (3) el formato de respuesta dicotómico (si-no; verdadero-falso); (4) la dificultad de comprensión, la longitud, el tono en el que están formulados los ítems, y la

redacción de los ítems de forma vaga (Echeburúa et al., 2003; Finlay y Lyons, 2002; Lemos, 2005; Nunnally y Bernstein, 1995); o (5) los temas sensibles (Gove y Geerken, 1977).

Otro tipo de estudios han hallado correlación negativa entre la aquiescencia y determinadas variables, entre las que destacan (1) el bajo nivel cultural y la baja capacidad de cognición (Ayidiya y McClendon, 1990; Bachman y O'Malley, 1984; DeLamater y McKinney, 1982; Ferrando y Anguiano-Carrasco, 2012; Finlay y Lyons, 2002; Greenleaf, 1992; Heaven, 1983; Krosnick, Nayaran y Smith, 1996; McClendon, 1991; Meisenberg, Lawless, Lambert, y Newton, 2006; Mirowsky y Ross, 1991; Morales, 2006; Narayan y Krosnick, 1996; Schuman y Preser, 1981; Ware, 1978); (2) bajos ingresos (Ware, 1978); (3) edad (Vigil-Colet, Ruiz-Pamies, Anguiano-Carrasco y Lorenzo-Seva, 2012) y (4) presión del tiempo, distracción, extroversión (Knowles y Nathan, 1997; McGee, 1967).

Proposición 16:

Se pueden atribuir los siguientes sesgos de respuesta a escalas de autoinforme utilizadas en prevención de riesgos psicosociales

- a) Aquiescencia
- b) La deseabilidad social
- c) El autoengaño o la simulación
- d) Otros sesgos

3.3. CONTROL DE SESGOS

F. Jiménez y Sánchez (2004) aseguran que ningún test está exento de fuentes de error y aun conociendo la importancia de los sesgos de respuesta como fuente de amenazas a la validez en los autoinformes y a pesar de que toda una rama de la ciencia psicológica se dedica desarrollar técnicas destinadas detectarlos y minimizarlos, éstos son a menudo eludidos (Dupuis et al., 2015). Sirva de ejemplo ilustrativo la revisión efectuada en el epígrafe 2.4 de esta tesis, donde solamente dos cuestionarios BFPSL y la batería de MC Mutua, hacen referencia explícita a medidas orientadas al control de la deseabilidad social. Precisamente, esta última batería también añade el control del factor sinceridad (Guàrdia et al., 2008; Salas y Meliá, 2004). Otro punto a destacar por parte de varios revisores de estas medidas (CEOE, 2013; Guàrdia, 2008b; Moreno y Báez, 2010) radica en determinar la calidad de los cuestionarios en base a su fundamentación teórica y sus propiedades psicométricas de fiabilidad y validez, sobre todo ponen el punto de atención en la validez de constructo, cuestión que es del todo acertada, sin embargo, no hacen referencia a la importancia que implica para la validez de un cuestionario la ausencia de sesgos.

Proposición 17:

A pesar de existir técnicas destinadas a detectar y minimizar los sesgos de respuesta a menudo son eludidos.

Proposición 18:

A la hora de administrar test de evaluación de riesgos psicosociales no es frecuente realizar control de posibles sesgos o fuentes de error.

La presencia de sesgos en un test no implica que toda la escala en su conjunto esté sesgada (Delgado, 1988), ya que existen procedimientos para detectar y/o equilibrar estos reactivos (Barker et al., 2002; Echeburúa et al., 2003). En general, para controlar y minimizar los sesgos de respuesta existen dos direcciones: (i) tenerlos en cuenta a la hora del diseño del instrumento de medida utilizando escalas *ad hoc* (ej., escalas de sinceridad) o a través de la formulación de los ítems (ej., inversión de ítems o ítems de respuesta forzada), de modo que se interviene en la forma de obtención de los datos; y (ii) una vez obtenida la información, se emplean técnicas estadísticas de tratamiento de datos para detectar los posibles sesgos (ej., transformación ipsativa en test de respuesta forzada, Cheung y Chan, 2002). El control de sesgos es muy importante para la mejora de la validez de los autoinformes. Por ejemplo, en el reciente estudio de Anguiano-Carrasco et al. (2013) han demostrado que el control de DS contribuye a minimizar el falseamiento de las respuestas.

3.3.1. Formas de control de la deseabilidad social en medidas de autoinforme

El estudio típico para aportar pruebas, a favor o en contra, de la habilidad para distorsionar los resultados es realizar comparaciones entre un grupo de control frente a un grupo de sujetos instruidos para responder con honestidad y/o instruidos para distorsionar positiva o negativamente las respuestas. Los cambios significativos en los perfiles de medias entre los grupos aportarían evidencia de la existencia de manipulación (Jenkins, 1984; Salgado, 2005). En virtud de ello, son varias las investigaciones que hallando diferencias significativas al instruir a los encuestados para falsificar las respuestas (de manera positiva o negativa), de modo que el grupo que conscientemente falseaba las respuestas en el sentido de las instrucciones obtenía puntuaciones mayores en DS que el grupo de control (Bartlett y Richard, 1967; Cabañeros, García Cueto y Lozano, 2003; Hogan, 1972; Mikesell, Calhoun y Lottman, 1970). Con todo, existen diferentes alternativas para controlar la DS que se revisan a continuación.

3.3.1.1. Escalas específicas

Dentro de la vasta literatura científica para detectar y medir la DS se encuentran las escalas específicas de medida. Se aplican paralelamente con la escala de interés y se comprueba si existe asociación significativa entre ellas (cálculo de correlaciones y análisis factoriales), en ese caso se supone que opera la DS (Enríquez y Domínguez, 2010). Las escalas más destacables son: (i) la escala unidimensional de medida formulada por Crowne y Marlowe (1960) denominada la *Marlowe-Crowne Social Desirability Scale* (MCSDS), que después de cumplir más de 50 años continúa siendo la escala de medida más notoria (Ferrando y Chico, 2000; Gutiérrez, Sanz, Espinosa, Gesteira y García-Vera, 2016) y (ii) el Inventario Balanceado de Conductas Deseables (BIDR) diseñado por Paulhus (1991) el cual fue diseñado para medir los dos constructos de autoengaño (AE) y manejo de impresiones (MI). Coexisten con ellas otras escalas de interés (ver tabla 24).

Tabla 24: Escalas de DS
Fuente: Elaboración propia

ESCALAS DE DESEABILIDAD SOCIAL
Escalas L y K del Inventario Multifásico de Personalidad de Minnesota (MMPI: Hathaway y McKinley, 1951): <ul style="list-style-type: none">- La escala L (Lie) trataría de detectar la predisposición de ofrecer una imagen positiva- La escala K de ofrecer una imagen negativa de uno mismo
<i>Good Impression Scale</i> (Escala de buena impresión, Gough, 1952)
Escala de deseabilidad social (Edwards, 1957)
Escala de deseabilidad social incluida en el primer cuestionario de personalidad <i>Maudsley Personality Inventory</i> (MP, Eysenck, 1959)
Escala de sinceridad de Eysenck (Eysenck y Eysenck, 1964)
Escala de distorsión motivacional positiva del 16PF (Cattell, Eber y Tatsuoka, 1970)
La Respuesta Deseable en Actitudes y Opiniones (Schuessler, Hittle y Cardascia, 1978)
Escala de deseabilidad social (<i>Occupational Personality Questionnaire</i> , SHL, 1999)

El sentido de este tipo de escalas estriba en decidir la eliminación o no de las respuestas de los encuestados en las que se haya detectado una puntuación alta en DS. Por ejemplo, Eysenck y Eysenck (1968) y Michaelis y Eysenck (1971) recomiendan eliminar el 5% de la muestra que contenga las puntuaciones menores en sinceridad, para luego dividirla en dos secciones (i) sujetos con alta puntuación en sinceridad y (ii) sujetos con baja puntuación en sinceridad. Con los dos grupos formados se realizarán análisis por separado y la decisión acerca de tener en cuenta o no las puntuaciones, dependerá de los correlatos con otras variables de personalidad incluidas en el cuestionario (ej., neuroticismo, extroversión). También pueden corregirse mecánicamente las puntuaciones de los sujetos altos en DS, mediante fórmulas matemáticas como, por ejemplo, una puntuación alta dos unidades de desviación respecto a la media (Salgado, 2005).

Estas técnicas también tienen sus detractores pues cabría la posibilidad de estar eliminando sujetos sinceros. En base a estos argumentos, Morales (2006) se hace eco de tres posibles interpretaciones para las respuestas de un encuestado con puntuación alta en DS, en donde solo la primera sería realmente una tendencia de respuesta capaz de invalidar los resultados:

- a) Se trata de sujetos con una mala imagen propia, con cierta anormalidad psicológica o conductual (depende del instrumento empleado) pero que responden para dar una buena imagen de sí mismos; sus respuestas son objetivamente falsas e insinceras.
- b) Se trata de sujetos normales con ajuste psicológico pobre, con limitaciones que podemos considerar importantes, pero que no se conocen bien; sus respuestas son sinceras, pero objetivamente falsas.
- c) Se trata de sujetos normales psicológicamente, con buena imagen propia; sus respuestas son sinceras y reflejan lo que son realmente (p. 239).

3.3.1.2. Otros métodos de control

A parte de las escalas específicas para detectar la DS existen otras recomendaciones, que aunque no son 100% infalibles, pueden minimizar las respuestas socialmente deseables si se utiliza de forma combinada (Salgado, 2005):

- a. Utilizar ítems socialmente neutrales y sutiles en vez de ítems obvios (De las Cuevas y González de Rivera, 1992; Hough, Eaton, Dunnette, Kamp y McCloy, 1990), ya que son difíciles de detectar por el encuestado al disimular lo que pretenden medir (Barker et al., 2002; Morales, 2006).
- b. Utilizar test de respuesta forzada (medidas ipsativas) con similar formato de deseabilidad (De las Cuevas y González de Rivera, 1992; Hough et al., 1990; Olea, Abad y Barrada, 2010; Salgado, 2005; Zavala, 1965), evitando así, el formato tipo Likert, qué según algunos expertos fomenta la DS (Schulz y Brese, 2008).
 1. Parafrasear los ítems en tercera persona para minimizar el involucramiento (Enríquez y Domínguez, 2010).
 2. Garantizar y enfatizar la confidencialidad y el anonimato. Cuanto más fácil sea identificar a los encuestados más propensos serán éstos a dar respuestas deseables (Lajunen y Summala, 2003; Morales, 2006; Paulhus y Reid, 1991; Podsakoff, MacKenzie, Lee y Podsakoff, 2003).
- c. La administración de los test vía internet. Las investigaciones revelan que las respuestas contienen menos DS porque la separación física entre la persona que aplica el test y el encuestado favorece el anonimato (Risko, Quilty y Oakman, 2006).
- d. Solicitar directamente sinceridad en las respuestas (Gordon, 1987; Podsakoff et al., 2003).
- e. Informar de consecuencias negativas para los distorsionadores de las respuestas (Converse et al., 2008; Hough et al., 1990; Pace y Borman, 2006; Salgado, 2005).
- f. Crear baremos (Salgado, 2005).

Existen estrategias de control de DS más complejas como (a) técnicas de análisis factorial para conseguir un factor específico de DS (ej., Ferrando, 2005; Ferrando y Anguiano-Carrasco, 2009, 2010; Ferrando et al., 2009; Neill y Jackson, 1970; Paulhus, 1981); (b) la aleatorización de preguntas (ej., De Jong, Pieters y Fox, 2010); (c) experimentación con manipulación inducida de las pruebas (Holden, 2007, 2008; Jansen, König, Kleinmann y Melchers, 2012; Robson, Jones y Abraham, 2008; Salgado 2005); y (d) se incluye a modo de curiosidad el Método *Bogus Pipeline* desarrollado por E. Jones y Sigall (1971). Se trata de conectar al encuestado a una máquina con indicador fisiológico, con objeto de que tenga el convencimiento de que serán detectadas sus creencias auténticas. A parte de su alto coste y de la complejidad del método, su gran limitación consiste en que rebasa las competencias profesionales de todo técnico de prevención de riesgos laborales y por lo tanto no es un método a tener en cuenta.

3.3.2. Formas de control de aquiescencia en medidas de autoinforme

3.3.2.1. Escalas específicas

Este tipo de escalas permiten controlar la aquiescencia como estilo de respuesta. Fueron muy usadas en la década de los 60 para luego caer en descrédito debido a sus bajas propiedades psicométricas, por lo que no son recomendadas para investigar tendencias generales de aquiescencia (Morales, 2006; Paulhus, 1991; Robins, Fraley y Kruger, 2007). Las escalas más populares han sido la escala de aquiescencia social de Bass (1956), la *Agreement Response*

Scale (ARS) de Couch y Keniston (1960), la escala de aquiescencia de Jackson y Messick (1961).

3.3.2.2. Otros métodos de control

3.3.2.2.1. Escalas balanceadas

Una de las medidas de control de aquiescencia más recomendada por los expertos es la construcción de escalas equilibradas o balanceo de ítems en contraposición a las escalas unidireccionales que son las que contienen ítems en una sola dirección (Ferrando et al., 2011; Likert, 1932; Morales, 2006, 2010; Nunnally y Bernstein, 1995; Oliver, Galiana, Sancho y Tomás, 2011; Prewett, Tett y Christiansen, 2013; Robins et al., 2007; Spector, 1992; Stricker, 2013; Tomás, Galiana, Hontangas, Oliver y Sancho, 2013). Esto es, aproximadamente el mismo número de ítems positivos y negativos, de manera que la mitad de la escala contenga ítems en dirección al constructo a medir (a mayor nivel en la dimensión, mayor puntuación en el ítem), y la otra mitad, en el sentido opuesto (Abad et al., 2006; Barker et al., 2002; Echeburúa et al., 2003; Ferrando y Lorenzo-Seva, 2010; Likert, 1932; Lorenzo-Seva y Ferrando, 2009; Paulhus, 1991). Es preciso aclarar que no se deben confundir los ítems negativos, con ítems redactados con partículas negativas (no, ninguno, nunca). No son formatos equivalentes (Morales, 2006). Tomás et al. (2012) para evitar la confusión, utilizan el término *ítem invertido* para referirse a los reactivos en sentido opuesto al rasgo e *ítem en negativo* para referirse a los que contienen en la redacción partículas negativas. Estos últimos son poco aconsejables ya que acentúan la ambigüedad del reactivo y contienen pobres propiedades de psicométricas (Abad et al., 2006; Ferrando y Lorenzo-Seva, 2010; Herche y Engelland, 1996).

Con todo, existe una corriente crítica con el procedimiento de inversión de ítems, aduciendo que disminuye la dimensionalidad (Harvey, Billings y Nilan, 1985; Herche y Engelland, 1996; Schriesheim y Hill, 1981) o presentan varianza de método (Tomás et al., 2013; Zhang, Noor y Savalei, 2016; Zhang y Savalei, 2015), entre otros errores. Sin embargo, Morales (2006, 2011) en su extensa revisión sobre el tema, propone varias razones por las cuales esta técnica es recomendable: (1) porque se formulan y responden con mayor atención, y estimulan en menor medida la respuesta aquiescente; y (2) porque cuentan con la ventaja de permitir comprobar la correlación entre los ítems invertidos y de ese modo se convierten en una medida de aquiescencia. El procedimiento, según Morales (2006), consiste en valorar con un punto cada respuesta positiva y con cero puntos las negativas, si la escala es dicotómica, si hay más de dos categorías de respuesta el máximo acuerdo obtendría la máxima puntuación. De modo se formarían dos grupos sobre los que se podrían aplicarse dos análisis: fiabilidad y el coeficiente de correlación entre los dos grupos de ítems (positivos e invertidos). DeVellis (1991) también indica que la correlación r entre los ítems en dirección al constructo a medir y los ítems en dirección opuesta, es la opción más factible para medir la aquiescencia, siendo $r \geq .50$.

3.3.2.2.2. Escalas con pares de ítems de significado opuesto

Una técnica de corrección menos utilizada y con falta de datos empíricos que lo evalúen, es la propuesta por Finlay y Lyons (2002) que consiste en utilizar pares de ítems con

significados lógicamente opuestos “(ej., ¿Estás feliz?, ¿estás triste?)” (p. 5). Morales (2006) también hace referencia a este método, y según explica, es considerado como variante de las escalas balanceadas y que requiere que no sólo contenga ítems positivos y negativos, sino que además “la escala esté compuesta de pares de ítems estricta y lógicamente opuestos entre sí, como *estudiar es entretenido* y *estudiar es aburrido* (evitando en lo posible expresiones como *no* y *nunca* para que no haya confusiones en las respuestas)” (p. 193). En estos casos los pares de ítems deben de ser opuestos y contradictorios a la vez, así la suma de respuestas contradictorias será la puntuación de aquiescencia. Cada vez que se responde estando de acuerdo con enunciados que se contradicen se marcaría un punto en aquiescencia, de modo que la puntuación máxima en dicho sesgo será igual al número de pares de ítems opuestos y contradictorios.

3.3.2.2.3. *Escalas con test ipsativos*

Otras soluciones aportadas para el control de la respuesta aquiescente es la utilización de formatos de test de respuesta forzada, también denominados test ipsativos (Calderón y Ximénez, 2014; Morales, 2006). Según Ray (1990), en este tipo de escalas la aquiescencia se muestra en la tendencia a escoger el primer enunciado. Es importante destacar que con este tipo de ítems no es aconsejable la utilización del análisis factorial convencional para la evaluación psicométrica, pues conduce a conclusiones erróneas (Calderón y Ximénez, 2014; Frutos, 2013). el sujeto debe ordenar distintos enunciados de acuerdo a cómo describen su comportamiento, respondiendo a todas las alternativas del bloque

3.3.2.2.4. *Otras recomendaciones para evitar la aquiescencia en los autoinformes*

Diversos autores entre los que se encuentra Morales (2006) y Nunnally y Bernstein (1995), opinan que si se redactan los instrumentos de medida correctamente el impacto de la aquiescencia sería insignificante. Otras posibles directrices a seguir serían (a) incluir en el cuestionario ítems manifiestamente falsos donde la respuesta debería ser no (Allen, 1966, citado por Morales, 2006); (b) incluir ítems sin sentido del estilo “¿Generalmente nieva aquí en verano? y ¿Sabes cómo pilotar un avión?” (Finlay y Lyons, 2002, p. 5); (c) confeccionar los instrumentos de medida pensando en la población objeto de estudio, pues como se ha mencionado en párrafos anteriores, se ha demostrado que determinados sujetos son más susceptibles a proporcionar respuestas aquiescentes (ej., mayores de 50 años, nivel cultural, educativo e intelectual bajo, escasos recursos económicos); (d) cuidar la redacción del cuestionario evitando enunciados complicados o demasiado vagos. Lo adecuado es incluir ítems de complejidad media y eludir ítems demasiado largos y con formato de respuesta dicotómico (Echeburúa et al., 2003; Finlay y Lyons, 2002; Lemos, 2005; Nunnally y Bernstein, 1995; Schuman y Presser, 1981); (e) evitar las escalas con muchos ítems, ya que propician la respuesta aquiescente por fatiga (Clancy y Wachslar, 1971, citado por Ferrando y Anguiano-Carrasco, 2012); y (f) utilizar procedimientos específicos basados en análisis factorial (Ferrando et al., 2009; Lorenzo-Seva y Ferrando, 2009, 2010).

3.4. SESGO DE VARIANZA COMÚN DEL MÉTODO

El sesgo de varianza común es habitual en estudios psicosociales y concretamente los cuestionarios de autoinforme son susceptibles a este sesgo (Jex, Crossley, 2005; Podsakoff et al., 2003; Razavi, 2001). Ocurre sobre todo cuando se utiliza una misma herramienta de medida (ej., cuestionarios) para evaluar distintas variables (Howard, 1994; Podsakoff y Organ 1986; Spector, 1994). La varianza común del método (VCM) se refiere a la varianza originada por el método de medida utilizado y no por la varianza cuyo origen son los propios factores o constructos observados. Littlewood (2006) recoge un completo listado de las fuentes de VCM propuestas por Podsakoff et al. (2003). Algunas de las distintas y numerosas situaciones que originarían la covarianza errónea se presentan en la tabla 25.

Tabla 25: Fuentes de Varianza Común de Método
Fuente: Elaboración propia

FUENTES DE VARIANZA COMÚN DE MÉTODO
Utilizar para todo el cuestionario el mismo tipo de escala (ej., escala Likert).
Medir distintos constructos con la misma herramienta.
Obtener las observaciones sobre variables criterio y predictoras del mismo informante.
La tendencia de los informantes de mantener consistencia en sus respuestas.
La tendencia de responder de manera socialmente deseable.
La tendencia de los encuestados de responder aquiescentemente.
El estado de ánimo (afectividad positiva-afectividad negativa), la tendencia del evaluado en percibir su entorno y a sí mismo en términos positivos o negativos.
Las características de los ítems según su redacción, por ejemplo, ítems redactados con DS, con sugerencia sobre la respuesta, ambiguos o confusos, con inclusión de expresiones positivas o negativas.
El orden de las variables independientes y dependientes puede influir en el participante al realizar asociaciones causales.
La longitud de la escala.

En todo caso, no tener en cuenta la VCM provocaría tener que enfrentarse a una posible fuente de error de medición, invalidando o distorsionando la validez de las relaciones entre las variables, es decir, las estimaciones y conclusiones realizadas sin contemplar la VCM serían potencialmente incorrectas a consecuencia de la inexactitud metodológica. En este sentido, Podsakoff et al. (2003) proponen varios procedimientos estadísticos para comprobar si la varianza se debe a los constructos que se están midiendo, como es lo esperable, o si por lo contrario es consecuencia del método de medida empleado, entre ellos destacan (1) el test de factor único de Harman; (2) el control de los efectos de factor latente (medido y no medido); (3) la correlación parcial (separando la DS y la afectividad; variables marca); (4) el uso de múltiples factores; (5) el modelo de correlaciones únicas; y (6) el modelo del producto directo.



PARTE 2: PARTE EMPÍRICA





PRIMER ESTUDIO





4.INVESTIGACION CUALITATIVA: METODO DELPHI

4.1. INTRODUCCIÓN

En la revisión de la literatura sobre la evaluación de riesgos psicosociales de origen laboral, ha quedado patente que son muchas las voces que cuestionan la validez de las medidas de autoinforme como herramienta de obtención de datos. El origen subjetivo de la información, la controversia científica de criterios acerca del instrumento de evaluación adecuado, el exiguo control de sesgos, la falta de protocolo oficial para la evaluación de dichos riesgos, entre otros motivos, han contribuido en gran medida a convertir la medición de los riesgos psicosociales en un maremagno. A pesar de su importancia, apenas existen trabajos científicos que examinen específica y detalladamente la validez de las medidas de auto-informe en el ámbito de la Psicología Aplicada, y en particular desde una aproximación cualitativa. En consecuencia, este estudio tiene como objetivo principal explorar las principales debilidades y fortalezas del uso de cuestionarios en la medición de riesgos psicosociales en el trabajo.

Analizar toda la diversidad de variables que puedan amenazar la validez en las mediciones de riesgos psicosociales conlleva una limitación tanto temporal como de recursos y, con objeto de ir desenredando los entresijos que rodean este tipo de evaluaciones y seguir avanzando en la investigación, se ha decidido someter al juicio de expertos las principales conclusiones que a modo de proposiciones se han ido subrayando de la revisión del marco teórico. Se pretende conseguir con ello, una composición de lugar desde el punto de vista de los profesionales implicados que permita, a su vez, identificar y definir un posible protocolo de acciones encaminadas a subsanar dichos problemas de validez.

Para ello, se ha utilizado la técnica Delphi, un procedimiento riguroso que permite, por un lado, recoger las valoraciones de los especialistas con el mayor grado de consenso posible y, por el otro lado, obtener aportaciones interesantes que contribuyan a la resolución de la problemática planteada. Si el juicio de un sólo especialista aporta valor, mayor credibilidad y confianza se obtendrá del consenso de un grupo de acreditados expertos (Miner, 1984). En este sentido, se ha confirmado ampliamente que el Delphi es una herramienta muy eficaz para alcanzar la concordancia de opiniones en situaciones de incertidumbre o de falta de evidencia empírica (Gil-Gómez y Pascual-Ezama, 2012; Linstone y Turoff, 1975; Pill, 1971).

4.1.1.Antecedentes del método Delphi

La técnica se gestó en California en la década de los 50 en el *think tank* (laboratorio de ideas) de la U.S. RAND (*Research and Development Corporation*). La RAND comenzó como institución subvencionada por las Fuerzas Aéreas de los Estados Unidos para dedicarse a investigaciones de seguridad nacional. En plena guerra fría, con la amenaza de las bombas

nucleares rusas, era necesario anticiparse a los futuros avances tecnológicos para la defensa del país y los métodos de predicción existentes no eran adecuados por la falta de datos históricos sobre el tema. Indagando sobre nuevos métodos de pronóstico nació el Proyecto Delphi (Tolon, 2012). Fue en 1953, de la mano de los matemáticos Norman Dalkey y Olaf Helmer, cuando se desarrolla por primera vez la técnica y se sientan las bases del método Delphi (Dalkey y Helmer, 1963). Para ello, se basaron en dos estudios anteriores:

- a) Experimento orquestado en la RAND, en el año 1948, donde un grupo de estadísticos pretendían mejorar las predicciones de los resultados de las carreras de caballos. Encontraron que cuando las apuestas estaban basadas en la opinión consensuada de un conjunto de *hadicappers* (expertos profesionales en valorar y establecer las escalas de pesos de los purasangres en competiciones) se tenían menos pérdidas en comparación con las apuestas que seguían el criterio de los expertos por separado (Gordon, 1994, Quade, 1967; Milliken, 1979).
- b) Experimento del filósofo Abraham Kaplan, quien logró demostrar de forma experimental la superioridad de la previsión grupal frente a la individual (Kaplan, Skogstad y Girschick, 1949).

Perfeccionando las técnicas anteriores y siguiendo el antiguo dicho de “dos cabezas son mejor que una” (Dalkey, 1969, p. 6.), ambos científicos resumieron de manera sistemática el juicio intuitivo de una selección de expertos. Concretamente, encuestaron a siete especialistas, mediante cinco cuestionarios, para estimar el número de bombas necesarias que podría soportar el país en caso de una hipotética guerra nuclear hasta que quedasen reducidos todos los arsenales estadounidenses. La mejora de la técnica, tal y como señala Dalkey (1969), es que “introdujeron una característica adicional, llamada repetición con retroalimentación controlada” (p. 15).

Según Dalkey (1969, p. 6) el método consta de 3 características inherentes: (i) está amparado en el anonimato, (ii) favorece la retroalimentación reiterada y controlada, y (iii) proporciona una respuesta estadística del grupo. Con estos elementos se corregían las limitaciones de experimentos anteriores como las influencias de miembros dominantes, el pensamiento grupal o las críticas y rencores por las opiniones divergentes (Dalkey y Rourke, 1971; Gordon, 2003; Okoli y Pawlowski, 2004; Rowe y Wright, 1999; Rowe, Wright y McColl, 2005). Además, se eliminan los juicios basados “en quien grita más fuerte” y se reducen las inhibiciones de los miembros de menor reputación (Milliken, 1979, p. 31). Ese proceso de repetición y retroalimentación en aras a conseguir el consenso del grupo recibió el nombre de “Delphi” como tributo al célebre oráculo de Delfos de la antigua Grecia (Dalkey, 1969, p. 15), un santuario dedicado al dios Apolo donde los visitantes consultaban acerca de su futuro cuestiones de papel político, militar y social y que era el de mayor reputación por el nivel de acierto en sus predicciones. Aunque fue Kaplan quien primero acuñó este término (Kaplan et al., 1949).

Dada la naturaleza militar del experimento de 1953, la existencia del método se mantuvo en secreto hasta una década después, cuando se desclasificó. Fue entonces cuando el Delphi dejó de utilizarse sólo para investigaciones militares y proliferó en el ámbito civil dentro de las áreas económica, política y social. Alcanzó gran popularidad gracias al *best seller* titulado *Report on a Long-Range Forecasting Study* (Gordon y Helmer, 1964) en el que se publicaban futuros hallazgos científicos y tecnológicos más allá del siglo XX, partiendo de la consulta a 82

expertos, entre los que destacaban figuras de renombre como Isaac Asimov, Artur Clarke y Stephen Toulmin, entre otros (Gordon, 2009a).

A partir de ese momento se disparó su fama y fue imparable su expansión, tanto territorial como temáticamente. Especialmente se abrió camino en los gobiernos, en empresas privadas, en el mundo académico y educativo, en los “*think tanks*”, y muchos otros. (Lang, 2000). Tan sólo en 10 años la cifra de investigaciones que utilizaban la técnica alcanzó los cuatro dígitos (Linstone y Turoff, 1975). En una revisión efectuada en el año 2008 en la base de datos *Scopus* buscando los términos “*Delphi study*” (estudio Delphi), aparecieron más de 15000 artículos publicados en ese mismo año en revistas profesionales con revisión por pares (Gordon, 2009b). Búsquedas similares sobre tesis defendidas hasta el año 2004 (Landeta, 2006; Rieger, 1986) ponen en evidencia el interés del método Delphi por la comunidad científica y justifican su vigencia y validez como herramienta de investigación.

Se ha utilizado en la multitud de ramas del saber, como en las ciencias de la salud (Castro, et al., 2014; de Meyrick, 2003; M. Fernández et al. 2015; Paraskevas, y Saunders, 2012; C. Pérez et al., 2013; Powell, 2003; Varela-Ruiz., Díaz-Bravo y García-Durán, 2012; Yáñez y Cuadra, 2008); deporte (Blasco, López y Mengual, 2010; Calabuig y Crespo, 2009); ingenierías agrícola, forestal y civil (Fernández de Castro y López, 2013; Manoliadis y Tsolas, 2010; Moreno, Traba y Morales, 2008; Rodríguez e Ibiert, 2010) y, por supuesto, también en la ciencia que nos ocupa: las ciencias sociales. Así, ha abarcado campos como educación (Cabero e Infante, 2014; Reguant-Álvarez y Torrado-Fonseca, 2016; Riaño y Palomino, 2015; Romero, Cabero, Llorente y Vázquez-Martínez, 2012); empresa (Arregui, Vallejo y Villarreal, 1996; Kauko y Palmroos, 2014; Mitchell, 1991; Tolon, 2012; Zárraga, Molina y Corana, 2013); política (Ali, 2005; Rosas, Sánchez y Chávez, 2012); psicología (Escobar-Pérez y Cuervo-Martínez, 2008; M. Fernández et al., 2015; Gil-Gómez y Pascual-Ezama, 2012; Hsu y Sandford, 2007; Landeta, 2006; Plans y León, 2003); o prevención de riesgos laborales (Boix, Gil y Rodrigo, 2008; J.M. Cortés, 2012; J.M. Cortés, Pellicer y Catalá, 2012; Morillas, Rubio-Romero y Fuertes, 2013; Rabada y Artazcoz, 2002; Reavley, Ross, Killackey y Jorm, 2012; Reavley, Ross, Martin, LaMontagne. y Jorm, 2014; University of Melbourne, 2013).

Con tanta profusión, el Delphi fue evolucionando y se ha modificado hasta dar lugar a una familia de técnicas inspiradas en el Delphi original, las más citadas son Delphi político, convencional y decisonal (Lang, 2000), conferencia Delphi, también denominado Delphi a tiempo real o *real-lucre delphi* (Linstone y Turoff, 1975, p.5), cara a cara (Mitchell, 1991), modificado EFTE (*Estimate, Feedback, Talk, Estimate*) (Nelms y Porter, 1991). Aunque existen otras muchas clasificaciones, algunas de ellas pueden consultarse en Brewer (2007), Hasson y Keeney (2011), Landeta (2002) o Mullen (2003).

4.1.2. Características principales

Un estudio Delphi trata de “un proceso sistemático e iterativo encaminado hacia la obtención de las opiniones, y si es posible del consenso, de un grupo de expertos” (Landeta, 2002, p. 32) y está basado en las siguientes características fundamentales: el anonimato, la iteración con retroalimentación controlada y respuesta estadística del grupo (Lang, 2000; Rowe y Wright, 2001; Rowe, Wright y Bolger 1991).

- a) Anonimato: es el elemento principal, pues no hay comunicación cara a cara, lo que permite a los individuos poder cambiar de opinión sin temor a la pérdida de imagen ante el grupo y poder aportar ideas sin recibir críticas ni dejarse influenciar por los juicios vertidos por los miembros dominantes o con más reputación. Con el anonimato se consigue minimizar la inhibición de respuesta ya que ninguno de los participantes conoce la identidad de las respuestas de los otros.
- b) Iteración: los expertos deben emitir su opinión en más de una ocasión. Son sucesivas rondas del proceso con tendencia a la convergencia de opiniones donde los participantes pueden reflexionar y reconsiderar sus planteamientos iniciales.
- c) Retroalimentación controlada: se reenvían a los participantes los resultados grupales de cada ronda, con sugerencias o razonamientos de otros miembros. El control de la comunicación por parte del investigador evita la iteración directa entre los expertos, lo cual reduce el efecto ruido (Dalkey, 1972), al evitar información irrelevante o errónea.
- d) Respuesta estadística del grupo: aparte de garantizar que estén incluidas todas las aportaciones de los miembros del grupo, ayuda a reducir la presión hacia la conformidad; al utilizar la mediana de las respuestas individuales se minimiza la necesidad de llegar a un consenso, el rango intercuartílico (RIR) de las estimaciones indicará el nivel de consenso o de dispersión.

Siguiendo estos fundamentos, el método Delphi se basa en recoger, mediante el envío sucesivo de encuestas anónimas, los juicios individuales de los expertos seleccionados con anterioridad. Cada fase de envío también es denominada “ronda”, “ola”, “circulación” o “vuelta”. Con los datos obtenidos en cada encuesta se efectúa un análisis estadístico que es puesto en conocimiento de los participantes en el siguiente envío de cuestionario, a los que se les solicita que, una vez conocida la distribución y dispersión de las respuestas, puedan ratificar o modificar su opinión inicial si así lo estiman apropiado. El objetivo de las iteraciones es conseguir la convergencia de opinión; o dicho de otro modo, la disminución de la dispersión de las respuestas.

Respetando estos cuatro elementos indispensables y calificados como universales (Landeta, 2002), la técnica puede ajustarse dependiendo del tema o de la información solicitada. El tamaño de la muestra, el número de rondas, la selección de expertos, o el tratamiento estadístico, son algunos de los elementos que han sido modificados por los investigadores según los objetivos a alcanzar (Rowe y Wright, 1999; 2001; Rosas et al., 2012; Woudenberg, 1991). Esta versatilidad de la técnica es una de sus mayores ventajas ya que, sin perder rigor, permite adaptarse a las condiciones del estudio que se lleve a cabo.

4.2. JUSTIFICACIÓN DEL PROBLEMA Y OBJETIVOS DE LA INVESTIGACIÓN

Tal y como se ha planteado a lo largo de esta tesis, la evaluación de riesgos psicosociales no carece de polémica, bien sea por la novedad de estos riesgos, por falta de legislación específica, por la escasa cultura preventiva, por su intangibilidad o por falta de formación e información, entre otras muchas circunstancias. Lo cierto es que desde un punto de vista práctico y preventivo existe un acusado ninguneo de estos riesgos, considerados desde hace tiempo como *la asignatura pendiente* de las agendas políticas y empresariales (Cifre et al, 2011; Fernández, R., 2010; M. García y Olarte, 2010; S. Jiménez, 2014; Leka y Jain, 2010; Merino, 2013; Cifre et al., 2009). Por lo grave de sus consecuencias, debería ser urgente atajar este

hecho, pues día a día está repercutiendo en la salud de los trabajadores y, a su vez, en la economía de las empresas. En la humilde tentativa de querer ayudar a enmendar este desaguisado, se ha creído que lo más eficaz sería comenzar por preguntar directamente a los expertos, tanto profesionales como académicos, acerca de cuáles son las verdaderas causas que, según su experiencia, están entorpeciendo las evaluaciones psicosociales. Los resultados de esta consulta contribuirían a detectar y priorizar los puntos clave sobre los que hay que trabajar para minimizar la problemática en la que se ha convertido evaluar riesgos psicosociales de origen laboral. Para dotar de peso y credibilidad al juicio de los expertos se ha optado por llevar a cabo la consulta a través del mencionado método Delphi.

El método Delphi consiste en “una opinión informada de personas con trayectoria en el tema, que son reconocidas por otros como expertos cualificados en éste, y que pueden dar información, evidencia, juicios y valoraciones” (Escobar-Pérez y Cuervo-Martínez, 2008, p. 29). En sus orígenes, tenía un enfoque puramente predictivo de escenarios futuros, pero ha pasado a convertirse en una herramienta esencial para establecer soluciones a problemas específicos y/o priorizar alternativas prácticas, lo que convierte al Delphi en una herramienta para la investigación y el aprendizaje (Gupta y Clarke, 1996; Murry y Hammons, 1995).

El estudio Delphi “es considerada la más destacada de las metodologías de consenso” (citado por Lang, 2000, p. 4), y dada la divergencia de opiniones acerca de la evaluación de riesgos psicosociales, junto con su carácter multidisciplinar, se ha creído acertado utilizar el Delphi con objeto de conseguir una visión experta sobre los principales retos a los que se enfrentan los académicos y profesionales implicados en la evaluación de riesgos psicosociales. La finalidad última es tratar de llegar a un conocimiento aumentado, gracias al consenso de opiniones basadas en la experiencia y sabiduría de expertos de distintos ámbitos. El punto de vista colectivo ofrecerá mayor confianza, “el resultado del comportamiento conjunto de un grupo de individuos es superior, bajo determinadas condiciones, a la suma de los resultados individuales” (Landeta, 2002, p. 23).

Por ende, es una técnica idónea para conseguir aportaciones valiosas a la complejidad del problema que se cuestiona. Tal escenario de investigación sólo es posible recabando las opiniones y valoraciones de los profesionales implicados. Como indican Gordon (1994), Strauss y Zeigler (1975) o Turoff (1970), a menudo el *quid* de la cuestión no reside en el consenso en sí, sino en la identificación de los juicios divergentes.

La decisión de su uso ha sido sustentada por ser un método empíricamente fiable de consenso en multitud de disciplinas (Brill Bishop, y Walker, 2006; Murry y Hammons, 1995) que, además, ha arrojado resultados de confianza en infinidad de estudios como, por ejemplo, los de Rohrbaugh (1979, citado por Rowe y Wright, 1999), donde se evidenció la supremacía de esta técnica comparada con otros métodos de decisión grupal; o los de Rowe y Wright (1999), que en su extensa revisión, dejan constancia de la efectividad y fiabilidad del método en variedad de estudios que utilizaron el método Delphi. Dalkey y Helmer (1963) apuntan que puede reunir “el consenso más fiable de opinión de un grupo de expertos” (p. 458). Otros motivos que sirvieron de aval en la elección de esta técnica han sido que, a diferencia de otras, permite la recogida de valoraciones y opiniones de expertos de un amplio ámbito geográfico y posibilita la comunicación a través del correo electrónico, lo que facilita la confidencialidad y el intercambio de información acortando los tiempos.

Aunque desde la sociología se ha clasificado el Delphi como un instrumento cualitativo, dicha categorización es objeto de debate (Pérez-Andrés, 2000) y una gran mayoría lo define como una combinación de procesos cualitativos y cuantitativos (Bourgeois, Pugmire, Stevenson, Swanson y Swanson, 2000), pues esta herramienta permite cuantificar con rigor el componente subjetivo de la información recogida.

4.2.1.Objetivos

El principal objetivo de este primer estudio es de carácter exploratorio. Como se ha apuntado anteriormente, se pretende conseguir una visión experta sobre los problemas que entraña evaluar los riesgos psicosociales en el trabajo. Para ello se ha consultado mediante el método Delphi a un conjunto de expertos, quienes, desde su experiencia y conocimiento profesional, han dado su valoración sobre el conjunto de proposiciones derivadas de la revisión de la literatura, construyendo así un retrato de la realidad aplicada que rodea dichas evaluaciones.

Como objetivos específicos se pretende:

1. Desde el punto de vista experto, elaborar una composición de lugar de la problemática que envuelve las evaluaciones de riesgos psicosociales en el trabajo.
2. Explorar las debilidades y fortalezas de dichas evaluaciones cuando son medidas con autoinformes.
3. Extraer directrices prácticas que consigan minimizar los puntos débiles de las evaluaciones psicosociales y puedan servir de orientación para el ejercicio profesional.

4.3. METODOLOGÍA DE ESTUDIO

Como en todo método, en el Delphi se siguen una serie de fases sucesivas (ver figura 7). En el presente estudio, una vez definido el problema, y siguiendo las pautas indicadas en la literatura, se procedió a configurar el panel de expertos y a confeccionar el cuestionario, se diseñó un Delphi de dos rondas con sus consiguientes explotaciones de resultados, para finalizar con el apartado de discusión de los resultados y conclusiones.

Figura 7: Fases del método Delphi
Fuente: Elaboración propia



4.3.1. Formulación del problema

Para proyectar a los expertos en la materia del problema de investigación, se realizó una revisión teórica del estado de la cuestión, donde fueron consultadas numerosas fuentes especializadas, como libros, revistas, tesis doctorales, legislación, recomendaciones del INSHT y de la ITSS, además de explorar diversas bases de datos bibliográficas: *Dialnet*, *Medline*, *ProQuest*, *PsycInfo*, *Science Direct* *Scifinder* y *SCOPUS*.

Efectuada la revisión bibliográfica, se han extraído las principales cuestiones problemáticas en cuanto a la validez de las medidas de autoinforme en las evaluaciones de riesgos psicosociales, articulando los ejes de discusión que servirán de guía a la hora de orientar las variables de estudio. Dichas cuestiones se han presentado a modo de proposiciones numeradas y, a partir de ellas, se ha diseñado la formulación del cuestionario con ítems cerrados y abiertos. Por lo tanto, el tipo de información solicitada a los expertos en el cuestionario está formada por dos tipos de preguntas diferentes. Por un lado, ítems cuantitativos tipo Likert (De Vaus, 2002) y, por otro lado, ítems cualitativos en forma de pregunta abierta. Para facilitar la comprensión del presente estudio, se enumeran a continuación las proposiciones recopiladas a lo largo de la revisión literaria que posteriormente fueron sometidas al juicio experto y que constituyen la base de la formulación de los ítems cuantitativos (tabla 26).

Tabla 26: Correspondencia entre ítems del cuestionario y proposiciones derivadas marco teórico
Fuente: Elaboración propia

PROPOSICIONES SUBRAYADAS DE LA REVISIÓN DEL MARCO TEÓRICO	
Ítem 1, corresponde a la proposición 1 de la revisión teórica:	
La mayoría de empresas concede más importancia a la evaluación de los riesgos físicos que a la evaluación de los riesgos psicosociales.	
Ítem 2, corresponde a la proposición 2 de la revisión teórica	
Son pocas las empresas que cuentan con procedimientos para gestionar riesgos psicosociales tales como estrés, <i>mobbing</i> , <i>burnout</i> , violencia, etc.	
Ítem 3, corresponde a la proposición 3 de la revisión teórica:	
A menudo se confunde lo que es un método de evaluación de riesgos (método científico), con el hecho de administrar cuestionarios que evalúan riesgos (técnica de recogida de datos).	
Ítem 4, corresponde a la proposición 4 de la revisión teórica	
La evaluación de riesgos implica más procesos que valorar los riesgos mediante un cuestionario.	
Ítem 5, corresponde a la proposición 5 de la revisión teórica	
Para valorar los riesgos psicosociales es esencial la autopercepción del trabajador.	
Ítem 6, corresponde a la proposición 6 de la revisión teórica	
Para que los datos recogidos sean lo más objetivos posible, la herramienta a utilizar para la recogida de datos debería contener criterios que cumplan las propiedades de fiabilidad y validez.	
Ítem 7, corresponde a la proposición 6 de la revisión teórica	
Para que los datos recogidos sean lo más objetivos posible, la herramienta a utilizar para la recogida de datos debería contener criterios que controlen los sesgos.	
Ítem 8, corresponde a la proposición 6 de la revisión teórica	
Para que los datos recogidos sean lo más objetivos posible, la herramienta a utilizar para la recogida de datos debería contener criterios de baremación y estandarización.	
Ítem 9, corresponde a la proposición 6 de la revisión teórica	
Para que los datos recogidos sean lo más objetivos posible, la herramienta a utilizar para la recogida de datos debería contener fundamentación teórica.	
Ítem 10. corresponde a la proposición 6 de la revisión teórica	
La herramienta de recogida de información debería contener otros criterios, no indicados en el cuestionario, para ser lo más objetivos posible.	
Ítem 11, corresponde a la proposición 7 de la revisión teórica	
Los cuestionarios de evaluación de riesgos psicosociales no deben incluir escalas cuya finalidad no sea detectar riesgos, como escalas de satisfacción o clima laboral.	
Ítem 12, corresponde a la proposición 14 de la revisión teórica	
Las medidas de autoinforme son susceptibles de ser falseadas.	
Ítem 13, corresponde a la proposición 9 de la revisión teórica	
A la hora de aplicar un test suele darse más importancia a las propiedades psicométricas de fiabilidad y validez que a la detección de sesgos.	
Ítem 14, corresponde a la proposición 16 de la revisión teórica	
La aquiescencia es uno de los sesgos que pueden atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales.	

*Continuación tabla 26***Ítem15, corresponde a la proposición 16 de la revisión teórica**

La deseabilidad social es uno de los sesgos que pueden atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales.

Ítem 16, corresponde a la proposición 16 de la revisión teórica

El autoengaño o la simulación son sesgos que pueden atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales.

Ítem 17, corresponde a la proposición 16 de la revisión teórica

Otros sesgos de respuesta no indicados en el cuestionario pueden atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales.

Ítem 18, corresponde a la proposición 11 de la revisión teórica

Un protocolo de administración donde, por ejemplo, los encargados de administrar los test lean las instrucciones pausada y claramente y traten de reducir la ansiedad de los evaluados, contribuiría a minimizar los posibles sesgos de respuesta.

Ítem 19, corresponde a la proposición 18 de la revisión teórica

A la hora de administrar test de evaluación de riesgos psicosociales no es frecuente realizar control de posibles sesgos o fuentes de error.

Ítem 20, corresponde a la proposición 12 de la revisión teórica

Los sesgos son un problema importante pues podrían contaminar los resultados de medida de los test, pudiendo incluso invalidar toda la evaluación.

Ítem 21, corresponde a la proposición 13 de la revisión teórica

Los sesgos de respuesta son posiblemente los de mayor incidencia en los cuestionarios de autoinforme.

Ítem 22, corresponde a la proposición 8 de la revisión teórica

Actualmente apenas existen herramientas para evaluar riesgos psicosociales que contengan control de sesgos.

Ítem 23, corresponde a la proposición 17 de la revisión teórica

A pesar de existir técnicas destinadas a detectar y minimizar los sesgos de respuesta a menudo son eludidos.

Ítem 24, corresponde a la proposición 15 de la revisión teórica

La percepción de incumplimiento de contrato psicológico, por parte del trabajador, podría influir en el falseamiento consciente de las respuestas de escalas utilizadas en la evaluación de riesgos psicosociales.

Ítem25, corresponde a la proposición 10 de la revisión teórica

Es necesario tener conocimientos de psicometría para interpretar los resultados de las escalas de autoinforme utilizadas para valorar los riesgos psicosociales.

4.3.2. Selección de expertos

La selección de expertos se efectuó desde un doble enfoque, cualitativo y cuantitativo. Desde el punto de vista cualitativo se realiza la selección según la facultad del sujeto de aportar información relevante a la investigación, basándose en sus conocimientos y experiencia, así como en la posición, producción científica, motivación y disposición a colaborar. En cuanto al enfoque cuantitativo se debe precisar el tamaño de la muestra, teniendo en cuenta los postulados derivados de la literatura, el tiempo y los recursos disponibles.

4.3.2.1. Enfoque cuantitativo: Número de rondas y tamaño de la muestra.

Según la literatura no existe un número de vueltas ideal, pudiendo variar según el objetivo del estudio y de los temas a tratar. Así, Woudenberg (1991) indica una variación entre dos y diez rondas. La determinación del número de iteraciones suele estar definido por la obtención de un grado de consenso y estabilidad de respuestas suficiente en las sucesivas vueltas (Dajani, Sincoff y Talley, 1979, 1980). Sin embargo, un número excesivo de rondas que intenten reducir la inestabilidad o la falta de convergencia de opiniones, dilatará desmesuradamente la duración del estudio, pudiendo fomentar el cansancio de los panelistas e incitarlos al abandono o a un consenso forzado (Gordon, 1994; Brill et al., 2006). Según estimaciones efectuadas por Bardecki (1984, citado por Landeta y Barrutia, 2011), el porcentaje de abandonos se sitúa entre un 20 y un 30%, y para salvar dicha falta de interés, son muchos los autores que abogan por utilizar dos o tres rondas, garantizando así la continuidad del panel (Brill et al. 2006; Landeta, 2006; Ortega, 2008; Rowe y Wright, 2001). De hecho, son pocos los estudios profesionales que utilizan más de dos vueltas (Landeta, 2002; Landeta, Matey, Ruiz y Galter, 2008).

Es preciso tener en cuenta que al trabajar con muestras reducidas, por pequeña que sea la tasa de abandono, podría distorsionar los resultados finales del estudio. Además, los mayores cambios en las respuestas Delphi ocurren en las dos primeras rondas (Dodge y Clark, 1977). Siguiendo las propuestas anteriores y a Landeta (2002), quien indica que lo más habitual es prefijar las rondas de antemano, se ha diseñado la presente investigación pensando en dos circulaciones, intentando mantener la motivación y participación de los panelistas al minimizar el excesivo lapso de tiempo que conlleva la gestión de cada cuestionario entre rondas.

El cuanto al número recomendable de sujetos a participar en un Delphi tampoco ha sido objeto de consenso (Hsu y Sandford, 2007; Rowe y Wright, 2001), encontrándose estudios que fijan distintos límites (Bourgeois et al., 2006; Brockhoff, 1975; Gordon, 1994; Landeta, 2002; Ludwig, 1997; Okoli y Pawlowski, 2004). En lo que sí parecen estar de acuerdo mayoritariamente es en utilizar muestras de pequeño tamaño. Se estima que el rango óptimo oscila entre 7 y 50 participantes (Astarriaga, 2000; Delbecq, Van de Ven y Gustafson, 1975; Landeta, 2002; J. Ruiz y Ispizua, 1989). De hecho, los primeros estudios de Dalkey mostraron que el error disminuía de manera significativa por cada participante añadido hasta llegar a 7, además, la curva de saturación a partir de 30 individuos no mejoraba de manera proporcional la calidad de los datos obtenidos. Por lo tanto, aumentar la muestra no compensa dado el incremento de tiempo y trabajo que conlleva. En consecuencia, utilizar muestras de más de 50 panelistas no es frecuente ni recomendable (Landeta, 2002). El Delphi no está destinado a producir resultados estadísticamente significativos (Gordon, 1994; Powell, 2003), es decir, no pretende ser representativo del ámbito profesional de los componentes que lo integran; sin embargo, por la relevancia de los encuestados, “los resultados representan, ni más ni menos, que la síntesis del pensamiento del grupo” (Gordon, 1994, p. 4).

En línea con estos argumentos, el número de participantes se ha fijado entre un mínimo de 8 y un máximo 30, contemplando la visión académica y la profesional. Para procurar dicha muestra será necesario enviar una solicitud de participación a un número muy superior de especialistas, utilizando el muestreo no probabilístico intencional descrito en R. Hernández et al. (2014), pues este tipo de muestras “son de gran valor, pues logran obtener los casos (i.e.,

personas, objetos, contextos, situaciones) que interesan al investigador y que llegan a ofrecer una gran riqueza para la recolección y el análisis de los datos” (p. 190).

4.3.2.2. Enfoque cualitativo: Criterio de Selección de expertos

El siguiente paso a abordar para la utilización del Método Delphi será la selección de expertos, garantizando su anonimato y determinar el tamaño de la muestra. En este trabajo, la selección se llevó a cabo entre dos ámbitos profesionales con experiencias potencialmente distintas para enriquecer el análisis. Por un lado, académicos del ámbito de la Psicología con actividad investigadora en riesgos psicosociales, y por otro lado, técnicos de prevención en activo, especializados en la prevención de riesgos psicosociales, ambos con conocimiento y experiencia en la temática, cumpliendo así el criterio de heterogeneidad. Cabe destacar que la expresión de “experto” es objeto de debate en la literatura, por lo que se han considerado dos de las interpretaciones más citadas, a saber: “el experto se debe definir como alguien que aporte inputs relevantes” (Pill, 1971, p. 60) y “aquel individuo cuya situación y recursos personales le posibiliten contribuir positivamente a la consecución del fin que ha motivado la iniciación del trabajo Delphi” (Landeta, 2002, p. 57).

En cuanto al criterio de selección de expertos, cabe decir que es uno de los elementos más importantes del método. Muchos autores, como Bjill (1992, citado por Lang, 2000), Landeta (2002), Nworie (2011), Okoli y Pawlowski (2004) o Powell (2003), lo consideran el punto crítico, pues la calidad de los resultados y la credibilidad del estudio dependerá en gran medida de una efectiva selección de los expertos. Por consiguiente, la elección tiene que ser cuidadosa. Del análisis de la literatura se concluye que no existe un criterio único, sino más bien una diversidad. Existen tanto estudios que no utilizan ningún tamiz, como los que utilizan criterios más elaborados como biogramas o autoevaluaciones de conocimientos sobre la temática a tratar (*self ratings*) (Brockhoff, 1975, Cabero y Llorente, 2013; Dalkey, Brown y Cochran, 1970; Gordon, 2009a). En el presente trabajo se ha utilizado una combinación de los dos criterios. Se ha seguido la propuesta de Cabero y Barroso (2013), debido a que el *self rating* se ha confirmado como un cometido que eleva el nivel de precisión del grupo y es considerado como una herramienta efectiva para medir el grado de experto (Rowe y Wright, 1996).

En primer lugar, se realizó una búsqueda de autores nacionales, principalmente psicólogos con publicaciones originales sobre el tema de estudio, en las bases de datos bibliográficas PsycInfo, Science Direct y Dialnet. Para la localización de técnicos de prevención profesionales en riesgos psicosociales, se utiliza la técnica de *Snowball sampling* (muestreo de bola de nieve) (R. Hernández et al., 2014) originalmente formulado por Coleman (1958) y Goodman (1961). Este procedimiento se inició entre los ponentes y profesionales asistentes a las “XXII Jornadas de Prevención de Riesgos Laborales” organizadas por la Sociedad Gallega de Prevención de Riesgos Laborales y la Universidad de Vigo en el mes de noviembre del año 2015. Con los resultados se obtuvo una lista de posibles participantes y se elaboró un biograma para cada uno de ellos, obteniendo la información necesaria a partir de búsquedas en Google sobre experiencia profesional, académica e investigadora, producción científica, participación en ponencias o reconocido prestigio en el sector al que pertenece (Escobar-Pérez y Cuervo-Martínez, 2008). Resultando así un listado de 94 especialistas a los cuales se solicitó su participación.

A los miembros que aceptaron colaborar en la investigación, como filtro adicional para determinar el nivel de experto, se les pidió que realizaran una autoevaluación sobre sus conocimientos en el ámbito de la evaluación de riesgos psicosociales. Para el cálculo del referido parámetro se ha utilizado una técnica en auge en los últimos años (Cabero y Llorente, 2013), el Coeficiente de Competencia de Experto (K) que está basado en la cualimetría, una metodología elaborada por el Comité Estatal para la Ciencia y la Técnica de Rusia en el año 1971 y que se emplea para determinar el coeficiente de competencia de un candidato como experto, siendo “1” el valor máximo (Fernández de Castro y López, 2013; L. García y Fernández, 2008; Gorina y Alonso, 2013; Zárraga et al., 2013). Se expresa formalmente del modo que sigue:

$$K = \frac{1}{2}(K_c + K_a)$$

Se trata de una autovaloración donde se indica el grado de conocimiento sobre el tema objeto de investigación y el grado de influencia de una serie de fuentes de argumentación que le permiten justificar dicho nivel (ver tabla 27). De acuerdo con los estudios de Cabero y Barroso (2013), dicha técnica “permite discriminar de forma más adecuada la selección de los expertos” (p. 34).

Según diversos autores citados por Zartha, Montes, Toro y Villada (2014), existe un consenso en calificar como (a) *competencia alta* los valores de *K* iguales o superiores a “.80”, (b) *competencia media* los valores inferiores a “.80” y superiores a “.70”, y (c) *competencia baja* los valores menores de “.70”.

Tabla 27: Valoraciones para el cálculo del Coeficiente de conocimiento de experto
Fuente: Elaboración propia adaptado de L. García y Fernández (2008) y de Cabero y Barroso (2013)

Kc	Es el <i>Coeficiente de conocimiento</i> que se posee sobre el tema planteado. Para su cálculo el experto asignará un valor que, a su juicio, represente lo más fielmente posible el grado de información que tiene sobre el asunto a tratar, dentro de una escala de 0 a 1. Donde 0 representa el desconocimiento total y 10 el pleno conocimiento en la materia, el resultado se multiplica por 0.1										
	0	1	2	3	4	5	6	7	8	9	10
Ka	Corresponde al <i>Coeficiente de argumentación</i> . Para hallar su valor el experto debe indicar el grado de influencia que considera que han tenido cada una de las seis fuentes indicadas en la tabla, en la fundamentación de sus criterios acerca de la materia objeto de debate. A cada grado marcado le corresponde la puntuación siguiente:										
	FUENTES			Grado de influencia de cada una de las fuentes en sus criterios							
				A (alto)	M (medio)	B (bajo)					
1º	Análisis teóricos realizados por usted			.30	.20	.10					
2º	Su experiencia en el tema			.50	.40	.20					
3º	Trabajos de autores nacionales consultados			.05	.05	.05					
4º	Trabajos de autores extranjeros consultados			.05	.05	.05					
5º	Su propio conocimiento del estado del problema en el extranjero			.05	.05	.05					
6º	Su formación			.05	.05	.05					

4.3.1. Control de errores y sesgos

Para el control de posibles errores o sesgos en el estudio se tomaron las siguientes estrategias: (1) seguimiento escrupuloso de los pasos de la metodología Delphi según la revisión de la literatura de referencia; (2) cada uno de los miembros del panel Delphi contó con la posibilidad de realizar aportaciones en cada apartado del cuestionario; (3) se ofreció a cada experto en la segunda ronda la retroalimentación estadística de las valoraciones del grupo y de las aportaciones adicionales, en base a la mediana de respuestas individuales, para reducir la presión hacia la conformidad y evitar anclajes de opiniones; (4) en todo momento se preservó el anonimato de los componentes del panel, sin posibilidad de que pudiesen ponerse en contacto para evitar el sesgo de efecto halo (Thorndike, 1920). Con el anonimato también se consigue minimizar el efecto espiral de silencio o la tendencia a enmudecer si se percibe que el resto del grupo opina de distinto modo para evitar el rechazo o el señalamiento (J. A. Anderson, 1996; K. Miller 2005; Noëlle-Neumann, 1995); (5) el *feedback* proporcionado a los expertos, referente a las aportaciones individuales de los panelistas, se presentó incluyendo la transcripción literal de los comentarios intentando así evitar sesgos de interpretación; (6) al utilizar el correo electrónico como medio de recogida y envío de información, estaban disponibles para su consulta todos los archivos enviados y recibidos; y (7) para evitar los sesgos de medición en el procesamiento de los datos, se revisaron concienzudamente tres veces cada uno de los cuestionarios recibidos.

4.3.2. Gestión 1ª ola

4.3.2.1. Confección y envío de cuestionario

Una vez resuelta la lista de panelistas, se confeccionó como instrumento de recolección de datos, garantizando el anonimato de cada participante, un cuestionario con ítems tanto cualitativos como cuantitativos. Los ítems cuantitativos utilizados son de escalamiento tipo Likert de 5 niveles, con los siguientes calificadores lingüísticos: Totalmente en desacuerdo (1), en desacuerdo (2), indeciso (3), de acuerdo (4) y totalmente de acuerdo (5). Se ha seleccionado este tipo de escala por ser la más recomendada por un gran número de expertos y por ser la preferida por los encuestados (Llaneza, 2013).

Atendiendo a esta distribución se formularon un total de 32 ítems cuidadosamente contruidos para evitar sesgos (25 ítems cerrados y 7 ítems abiertos). Fueron elaborados a partir de las principales cuestiones controvertidas halladas tras la revisión de la literatura científica y las guías de referencia de las principales instituciones públicas y privadas en el campo de la Prevención de Riesgos Laborales. Estos ítems se centran en el uso de cuestionarios en la evaluación de los riesgos psicosociales y, de acuerdo a su contenido, fueron divididos en 4 apartados:

- I. Afirmaciones generales (4 ítems cuantitativos: A1-A4), ej., *“la mayoría de empresas concede más importancia a la evaluación de los riesgos físicos que a la evaluación de los riesgos psicosociales”*.
- II. Diseño de instrumentos de medida (7 ítems cuantitativos + 1 cualitativo: A5-A11), ej., *“para valorar los riesgos psicosociales es esencial la autopercepción del trabajador”*.

- III. Aplicación (7 ítems cuantitativos + 1 ítem cualitativo: A12-A18), ej., *“las medidas de auto-informe son susceptibles de ser falseadas”*.
- IV. Corrección e interpretación de los resultados (7 ítems cuantitativos: A19-A25), ej., *“es necesario tener conocimientos de psicometría para interpretar los resultados de las escalas de auto-informe utilizadas para valorar los riesgos psicosociales”*.

Para completar la información de los apartados II y III, se formulan dos cuestiones semicerradas con la opción “otros” que corresponden a los ítems “10a” y “17a”. Dado que los ítems cualitativos son fundamentales para aprovechar el conocimiento y experiencia de los panelistas en la materia, al final de cada dimensión se solicita a los expertos que, en formato de texto libre, expresasen cualquier reflexión o propuesta que considerasen relevante (4 ítems cualitativos). Adicionalmente, se incluyó un apartado final de comentarios y/o sugerencias (1 ítem cualitativo). Los ítems sin respuesta se consideraron a efectos estadísticos como casos perdidos.

En el inicio del cuestionario se aportó información de los plazos de remisión y a su vez se incluyó una pequeña batería de preguntas de autoevaluación destinada a conocer el Coeficiente de competencia de experto (K), por ser ese el criterio seleccionado de *self rating*. Para crear el formulario se utilizó el programa *Adobe Acrobat DC* versión 2015. El cuestionario en concreto puede consultarse en el Anexo I.

Aprovechando las ventajas de las nuevas tecnologías, en fecha de 26 de noviembre de 2015, se envió el cuestionario a través del correo electrónico a cada uno de los 94 expertos primeramente seleccionados. En el cuerpo del correo se realizó un escrito de presentación, incluyendo la solicitud de colaboración, informando del objetivo de la investigación y de las instrucciones para rellenar el test. Asimismo, se indicó como primer plazo para recibir contestación el 30 de diciembre del mismo año. Llegada esa fecha se recibieron las respuestas de 16 expertos, y teniendo en cuenta que ese periodo abarcaba las fiestas navideñas, se creyó conveniente aumentarlo 20 días más. Se informó del nuevo calendario a todos los miembros contactados, y aprovechando la ampliación temporal, se incrementó la lista inicial con 23 nuevos expertos, siguiendo la misma técnica de muestreo. Terminado el plazo, el total de miembros dispuestos a colaborar fue de 20 especialistas. Las fechas de envío y recibo de los cuestionarios quedaron fijadas como se muestra en la tabla 28.

Tabla 28: Calendario de envío y recepción de cuestionarios
Fuente: Elaboración propia

RONDAS	FECHAS DE ENVIO	FECHA LIMITE DE ADMISION
1ª RONDA (1R)	26 de noviembre de 2015	20 de enero de 2016
2ª RONDA (2R)	3 de febrero de 2016	10 de marzo de 2016

Inicialmente, los plazos previstos para la recepción de cuestionarios en cada ronda fueron pensados para un periodo de 30 días, sin embargo, al coincidir éstos con periodos festivos (celebraciones navideñas y carnavales), se ampliaron los plazos con objeto de facilitar la obtención de respuestas. A pesar de ello el tiempo del proceso se encuentra dentro de los límites razonables (Delbecq et al., 1975; Landeta, 2002).

4.3.2.2. Valoraciones de expertos y tratamiento estadístico

Una vez recibidas las respuestas de los participantes, se confeccionó la lista definitiva de panelistas atendiendo a su grado de experto. El filtro a utilizar es el de mayor coeficiente de competencia (K), teniendo en cuenta que cuanto más se acerque “K” a uno, mayor será el grado de competencia. Existen estudios donde el punto de corte utilizado es ≥ 0.7 (Castillo, Díaz, Fariñas y Delgado, 2009), e incluso ≤ 0.7 (Riaño y Palomino, 2015); sin embargo, existe consenso generalizado de descartar a los candidatos con valores de $K \leq 0.8$ (Blasco et al., 2010; Cabero y Barroso, 2013; Cabero e Infante, 2014; Castro et al., 2014; L. García y Fernández, 2008; A. González, Guerrero y Maldonado, 2001; Mengual, 2011, Oñate, 2001, citado por Zarth et al., 2014). Cumpliendo esta última premisa, el panel de expertos quedó conformado por los 16 expertos con valores $K \geq .8$, prescindiendo de los 4 miembros restantes por no alcanzar dicho valor (tabla 29).

Tabla 29: Composición del panel de 16 expertos en materia de riesgos psicosociales
Fuente: Elaboración propia

FORMACION	PROFESION		
	Técnicos prevención	Académicos	Ambos
Prevención de riesgos laborales	6	-	-
Psicólogos	3	5	-
Ambos	-	1	1

El tratamiento de los datos siguió un proceso riguroso y meticuloso con el fin de conseguir interpretaciones exentas de toda subjetividad posible. Para ello, se llevaron a cabo distintos tipos de análisis estadísticos, entre ellos los que figuran en la tabla 30. Las herramientas utilizadas para tal fin fueron el procesador estadístico SPSS versión 17.0 y la hoja de cálculo Excel 2013.

Atendiendo a que el número de circulaciones fue fijado de antemano en dos rondas, no se atiende a los criterios de consenso o estabilidad como decisión de conclusión de las iteraciones. Sin embargo, dichos criterios fueron utilizados con fines exploratorios para conocer el grado de convergencia o divergencia de los panelistas y, por lo tanto, guiaron el análisis de la información aportada por los expertos. El investigador debe seguir un procedimiento disciplinado y suficientemente riguroso de manera que evite en la medida de lo posible las interpretaciones demasiado subjetivas de esos datos, pero manteniendo su sentido y sin perder información real (Landeta, 2002, p. 76).

4.3.2.3. Análisis estadístico 1ª ronda: Posición, tendencia central, consenso y dispersión

El tratamiento de los datos cuantitativos obtenidos de las opiniones de la primera ola, se realizó calculando la mediana (Me) como medida de tendencia central, los cuartiles como normas de posición (Q1, Q2, Q3), el rango intercuartílico relativo (RIR) como índice de consenso, y el rango intercuartílico (RI), la desviación típica (σ) y el coeficiente de variación (CV) como medidas de dispersión, por ser éstas las medidas utilizadas frecuentemente en este método para la presentación de resultados y para ofrecer *feedback* a los participantes (Landeta, 2002; I.M. Martínez, 2003; Von der Gracht, 2012).

Tabla 30: Estadísticos 1ª Ronda (1R)
Elaboración propia

Tipo de análisis	Estadísticos	Forma de cálculo	Criterios de análisis
Tendencia central	Me	$\frac{(X_{N/2} + X_{N/2 + 1})}{2}$	Me > 2 y < 4 ⇒ “indeterminados” (ni acuerdo ni desacuerdo)
			Siempre que menos de 1/3 de datos se encuentra fuera del rango donde se sitúa Me: Me ≥ 4 ⇒ “acuerdo a favor” Me ≤ 2 ⇒ “desacuerdo o acuerdo en contra”
Posición	Q ₁ , Q ₂ , Q ₃ Q ₂ = Me	Q _k = k (N/4) Q _k = Cuartil: 1, 2, 3 o 4	Los cuartiles dan información sobre la dispersión de las respuestas del grupo
Nivel de consenso y dispersión	RIR	RIR = (Q ₃ -Q ₁)/Me	RIR > 0.50 ⇒ falta de consenso RIR < 0.25 ⇒ consenso amplio RIR = 0 ⇒ consenso muy significativo
	RI	RI = (Q ₃ -Q ₁)	< RI, Q ⇒ mayor consenso
	Q	Q = (Q ₃ -Q ₁)/2	> RI, Q ⇒ mayor dispersión
	σ	$\sqrt{\frac{\sum x_i^2 \cdot n_i}{n} - \bar{x}^2}$	> σ ⇒ mayor dispersión
	CV	CV = (σ) / (\bar{x})	CV < 0.26 ⇒ muy baja variabilidad

Me = Mediana / Q_x = Cuartiles / Q₂ = Semi intercuartil / RI = Rango Intercuartílico / RIR = Rango Intercuartílico Relativo
σ = Desviación típica / CV = Coeficiente de Variación / \bar{x} = Media

La mediana corresponde al valor central de un número de observaciones ordenadas de mayor a menor; es decir, es el valor que divide la distribución en dos partes iguales porque deja el 50% de los datos a la derecha y a la izquierda. Es más robusta que la media, pues evita la distorsión de los valores extremos (Arregui et al., 1996; Behar y Grima, 2012; Canavos, 2003) y la más adecuada para escalas tipo Likert (Eckman, 1983; Hill y Fowles, 1975; Jacobs, 1996, citado por Hsu y Standford, 2007).

Los cuartiles son útiles para mostrar la distribución de las respuestas del grupo. Son los que dividen los datos ordenados en cuatro partes iguales, siendo el primer cuartil (Q₁) el valor que deja el 25% de observaciones a su izquierda, el segundo cuartil (Q₂) corresponde a la mediana, pues deja el 50% de datos a la derecha y el otro 50% a la izquierda, y por último el tercer cuartil (Q₃) es el valor que deja a su izquierda el 75% de las respuestas obtenidas. El rango intercuartílico relativo (RIR) es la medida utilizada como índice de consenso. Un RIR nulo sería RIR = 0, cuyo significado equivaldría a decir que al menos un 50% de los expertos respondieron coincidiendo con el mismo valor que la mediana, lo cual implica un nivel de consenso significativo.

Aunque se promueva el consenso, éste no es indispensable, pues no necesariamente tiene que obtenerse. Lo que sí es un objetivo es lograr una distribución lo menos dispersa posible y este hecho hace indispensable contar con índices que indiquen el grado de variación de las respuestas, permitiendo identificar las respuestas con el menor grado de dispersión posible. Para este análisis se han utilizado el rango intercuartílico, la desviación típica y el coeficiente de variación. El Rango intercuartílico (RI) es el resultado de la diferencia entre el tercer y el primer

cuartil ($Q3-Q1$) y permite el reconocimiento de variación en las opiniones de los expertos (Nick y Lee-Ross, 1998). La amplitud semicuartil (Q), es otra medida de dispersión que corresponde a la mitad del RI, indicando el grado de dispersión de los datos alrededor de la mediana. La desviación típica (σ) y el Coeficiente de variación (CV) permiten completar el análisis de dispersión de las respuestas. La desviación típica indica la desviación respecto a la media, cuanto mayor sea (σ), mayor será la dispersión de los valores respecto a la media aritmética y, por lo tanto, menos representativa será la media. El CV es una medida adimensional que sirve para comparar la heterogeneidad o dispersión de dos o más grupos y se calcula dividiendo la desviación típica (σ) entre la media aritmética (\bar{X}), y se multiplica por 100 para expresarlo en porcentaje.

4.3.2.4. Criterios de decisión de consenso

Según Luna, Infante y Martínez (2006) “el consenso se obtiene por un procedimiento matemático de agregación de juicios individuales” (p. 85, citado por Cabero e Infante, 2014) y para obtenerlo se necesita que las valoraciones de los expertos alcancen un grado suficiente de afinidad. En consecuencia, para un análisis preliminar de la concordancia de la opinión de expertos se han utilizado dos criterios principales de decisión, empleados en estudios Delphi por otros autores como, por ejemplo, Boix et al (2008):

- 1º. El porcentaje de acuerdo según la posición de la mediana como primera norma de discriminación de ítems no consensuados. Para ello, la escala de valoraciones se agrupa en tres intervalos con las siguientes puntuaciones [1,2], [3], [4,5] -en adelante, [C1-C2], [C3] y [C4-C5]-. De este modo, se considera que existe consenso cuando las puntuaciones de los panelistas que se encuentran fuera del rango donde se encuentra la mediana son menos de una tercera parte de los encuestados, es decir, un 33%. En ese caso es la posición de la mediana la que determina el acuerdo, se considera consenso en la categoría “acuerdo” cuando la mediana es ≥ 4 ; consenso en la categoría “desacuerdo” si la mediana es ≤ 2 ; los casos en los que la mediana es >2 y <4 se consideran “indeterminados”. Según lo expuesto no habrá convergencia de opiniones cuando más de $1/3$ de los datos se sitúan fuera del rango que determina la mediana (Boix, et al., 2008; C. Pérez et al., 2013; Seguí y Server, 2010).
- 2º. El segundo criterio de decisión tiene que ver con los resultados del rango intercuartílico relativo (RIR), donde $RIR > 0.50$ implica falta de consenso.

Para reforzar ambos criterios de decisión, se completó el tratamiento estadístico con la desviación típica y el coeficiente de variación, medidas que, aunque alternativas al RIR (Ferrán, 2001), informan de la baja variabilidad de respuesta. De este modo, quedaron fijados los parámetros que mostrarían las preguntas con resultados de no consenso e indeterminados (ver tabla 31). Estos últimos se propondrían posteriormente para su reconsideración en una segunda iteración.

Tabla 31: Criterios para considerar la falta de consenso en los ítems
Fuente: Elaboración propia

ORDEN	CRITERIOS DE DECISIÓN	RESULTADO
1º	$Me > 2 < 4$	Indeterminado: (ni acuerdo, ni desacuerdo)
2º	$RIR \geq 0.50$	No consenso
3º	% de respuestas fuera del rango donde se sitúa la Me $> 33\%$	No consenso
ORDEN	CRITERIOS DE REFUERZO	RESULTADO
4º	$\sigma \geq 1$	No consenso por exceso de variabilidad de respuesta
5º	$CV \geq 0.26$ (26%)	

Me = Mediana / RIR = Rango Intercuartílico Relativo / σ = Desviación típica / CV = Coeficiente de Variación

Para el resto ítems que no cumplan estos baremos se presume que existe consenso. Atendiendo a sus porcentajes de respuesta en los rangos [C1, C2] y [C4, C5], se realiza la graduación siguiendo premisas de otros estudios (Boix, et al., 2008; Landeta, 2002; C. Pérez et al., 2013; ver tabla 32).

Tabla 32: Grado de acuerdo alcanzado según la concordancia de opiniones
Fuente: Elaboración propia

	COINCIDENCIA PUNTUACIONES	RANGO [C1,C2]	RANGO [C4,C5]
RIR = 0	100%	Unanimidad en contra	Unanimidad a favor
RIR > 0 < 0.25	$\geq 80\% < 100\%$	Consenso amplio en contra	Consenso amplio a favor
RIR $\geq 0.25 < 0.5$	$< 80\% \geq 67\%*$	Acuerdo suficiente	Acuerdo suficiente
RIR ≥ 0.5	$< 67\%$	No consenso	No consenso

* El 67% corresponde a 2/3 de los participantes

Nota: No se incluye el valor C3 (3) pues corresponde a la opción de "ni acuerdo, ni desacuerdo"

4.3.3. Gestión 2ª ola

4.3.3.1. Confección y envío del cuestionario

En vista de los resultados que se alcanzaron en la ronda 1, el propósito de la segunda fue promover el consenso obteniendo una menor dispersión de respuestas (Astigarraga, 2006). Para ello, el 3 de febrero del 2016, se devolvió el cuestionario a los 16 panelistas participantes, informándoles de los resultados de las respuestas de todo el panel y resaltando los aspectos objeto de discrepancia, los indeterminados o dudosos, y aquellos que obtuvieron alta dispersión. Se les pidió que volviesen a reconsiderar y valorar estos ítems y aquellos otros que considerasen objeto de reflexión. Las respuestas de los ítems abiertos dieron lugar a nuevas cuestiones que se añadieron al cuestionario como ítems cuantitativos para su valoración por parte de los expertos. Por consiguiente, el cuestionario que se confeccionó en esta segunda circulación estaba dividido en tres partes diferenciadas:

- I. En la primera se incorporó una tabla con un resumen de los análisis estadísticos efectuados, mostrando las frecuencias relativas y las medidas de dispersión obtenidas en cada uno de los ítems. Se identificaron las preguntas que no obtuvieron convergencia de opinión, dejando las filas correspondientes en blanco y sombreando en verde el resto de los ítems que alcanzaron el consenso según los criterios establecidos.

- II. El segundo bloque siguió la misma estructura y los mismos ítems del primer cuestionario, a excepción de las preguntas abiertas, que fueron eliminadas. En este apartado cobra sentido el tratamiento estadístico efectuado, pues a la luz de sus resultados los expertos pudieron reconsiderar (o no) todas y cada una de las preguntas que considerasen oportunas en la búsqueda de consenso.
- III. Finalmente, se estructuró una tercera parte, donde se incluyeron a modo de sumario y de forma cerrada las respuestas de los expertos a los ítems cualitativos del primer cuestionario, de forma que todos pudiesen conocer las aportaciones de todos los panelistas y valorarlas según el formato Likert de 5 puntos. Asimismo, se incluyó una sección final para que los panelistas que así lo desearan pudieran completar o aclarar aquellos aspectos que fueran de su interés.

La participación en esta ronda resultó mejor de lo esperado al participar el 100% de los panelistas, lo que refuerza la teoría de que un menor número de rondas favorece una tasa de abandono menor (Brill et al., 2006), avalando la decisión de utilizar un Delphi a dos vueltas.

4.3.3.2. Análisis estadístico y criterios de consenso 2ª ronda.

Con los resultados obtenidos de esta segunda iteración se abordó el análisis de los ítems cuantitativos y cualitativos. Con referencia a los ítems cerrados se realizó un análisis estadístico final aunando los criterios y comparando los cuestionarios de la ronda 1 (R1) y de la ronda 2 (R2). En consecuencia, se efectuaron, en primer lugar, los mismos análisis estadísticos de posición, tendencia central, dispersión y criterios de consenso que se efectuaron en la primera ola y, en segundo término, se analizó la estabilidad entre rondas, cuyo estudio se recoge en epígrafes posteriores. Los ítems cualitativos se ordenaron atendiendo a la tasa de respuesta, en orden decreciente de mayor a menor.

4.3.4. Estabilidad entre rondas

La existencia de consenso, en cierto modo implica estabilidad; sin embargo, ésta no es sinónimo de consenso. Independientemente del grado de convergencia alcanzado, se entiende por estabilidad la no variación estadística de los juicios de los expertos (Landeta, 2002). Aunque alcanzar consenso es deseable, no se ha pretendido utilizar la metodología Delphi en persecución del acuerdo absoluto entre expertos, sino como herramienta que evidencie sus distintas posturas, convergencias y divergencias y que pongan al descubierto el escenario actual de la realidad que envuelve el proceso de evaluación de los riesgos psicosociales. De ese modo, un elemento discordante puede resultar de gran importancia informativa. Lo que convierte la exploración de la estabilidad en un análisis de gran interés pues “permite analizar las distintas posiciones dentro del grupo en referencia a cada una de las cuestiones planteadas” (Landeta, 2002, p. 91).

Con independencia de que se alcance un mayor consenso, lo esperado sería que, una vez conocidos los resultados del grupo, un alto porcentaje de expertos madurase sus respuestas modificándolas en la segunda ronda, y que en sucesivas circulaciones apenas variasen de opinión, observando mayores niveles de estabilidad (Rowe y Wright, 1999). Por ello, es utilizada muy a menudo como criterio de clausura de circulaciones.

El siguiente paso que se impone es la elección de métodos que midan con diferentes análisis, la variación de la respuesta de los expertos entre ronda y ronda, y los baremos de

referencia que prefijen niveles de estabilidad suficiente. Sin embargo, como ya ha sido mencionado, y siguiendo recomendaciones de otros autores como Landeta (1999), el número de vueltas fue determinado a priori en dos rondas, como manera de garantizar la calidad de resultados al intentar reducir las tasas de abandono y los tiempos de gestión. Decisión que hace irrelevante utilizar índices previos de estabilidad como referencia para la finalización de las circulaciones y hará que se interprete como resultado favorable la observación de reducción de los índices de estabilidad entre las respuestas de la segunda vuelta respecto a la primera, pues implicaría que el *feedback* ha surtido efecto. A pesar de ello se indicarán a modo meramente informativo los topes utilizados con más frecuencia en los estudios Delphi para la decisión de una estabilidad favorable. En consecuencia, se pretende analizar el criterio de estabilidad independientemente del consenso, y utilizarlo como indicativo del posicionamiento de grupo.

Lo más frecuente en los estudios Delphi, es ofrecer a los panelistas la posibilidad de modificar a partir de la segunda ronda, los reactivos que no han alcanzado convergencia o han resultado dudosos. Sin embargo, en el presente Delphi se posibilitado a los expertos la opción de poder variar todos y cada uno de los ítems cuantitativos, incluso los consensuados, si así lo creían adecuado. Esta decisión hace suponer que un gran número de panelistas modificará su posición. Al respecto, cabe destacar que la estabilidad puede entenderse de dos formas, estabilidad grupal y estabilidad individual (Schaffin y Talley, 1980; Landeta, 2002).

La estabilidad grupal mide la estabilidad de la respuesta conjunta del grupo cuantificando la variación de la distribución de las respuestas de una ronda en comparación con la anterior. Es una medida menos fuerte que la individual, pero de mayor relevancia, al tener en cuenta que lo que interesa es la calidad del grupo de expertos en su conjunto, pues es lo que otorgará fiabilidad al método Delphi (Dajani et al., 1979, 1980; Landeta et al., 2008; Scheibe, Skutsch y Schofer, 1975; von der Gracht, 2012).

Coexisten numerosas soluciones de estabilidad grupal, entre ellas la variación relativa de la mediana, variación de los coeficientes de dispersión (Q, RIR, CV), comparación de varianzas mediante un contraste de F de Snedecor, T de Student para muestras relacionadas, test Chi-cuadrado (χ^2) y prueba de Wilcoxon test como pruebas no paramétricas, índices de concordancia (coeficiente Kappa de Cohen) y muchas otras (I.M. Martínez, 2003). Los medidores elegidos para cuantificar el nivel de estabilidad grupal han sido la Variación del RIR, la prueba de Wilcoxon y el Coeficiente Kappa de Cohen. Dichas medidas y sus baremos se presentan en la tabla 33, y los baremos tienen que ser cumplidos por el 80% de las observaciones para considerar que existe estabilidad (Camisón, Camisón, Fabra, Florés y Puig, 2009).

Tabla 33: Estadísticos de medida de estabilidad grupal
Fuente: Elaboración propia

Tipo de análisis	Estadísticos	Criterios de análisis *	Estabilidad
Variación de medidas de dispersión	$VRIR = RIR_k - RIR_{k-1}$	$-.25 \leq VRIR \leq .25$	Amplia
	$VCV = VCV_k - VCV_{k-1}$	$v \geq .20 \leq .27$	Suficiente
		$v > 0 < .20$	Amplia
		$v = 0$	Absoluta
Test no paramétricos	Prueba de Wilcoxon	Significancia $p < .05$	Insuficiente (muestras diferentes)
Índice de concordancia	Coefficiente Kappa de Cohen	(Ver tabla 34)	

*: Para considerar la estabilidad como adecuada el 80% de los ítems deben de cumplir los criterios especificados

RIR = Rango intercuartílico relativo / $VRIR$ = variación RIR / K = n° de rondas / CV = Coeficiente de variación / v = Variación coeficiente de variación (VCV)

4.3.4.1. Variación de las medidas de dispersión

Las variaciones de los coeficientes de dispersión son utilizadas frecuentemente como medidas de estabilidad grupal (Cabero e Infante, 2014; Ortega, 2008). En palabras de I.M. Martínez (2003), “dado el coeficiente de una ronda y la anterior, si la diferencia entre ambos es menor que un nivel establecido previamente se considerara aceptable la estabilidad de las respuestas” (p. 456). En este estudio se analizó la variación de los índices de dispersión del rango intercuartílico relativo (RIR) y del coeficiente de variación (CV).

a. La variación del rango intercuartílico relativo entre las dos rondas ($VRIR$).

Se obtiene calculando la diferencia entre el RIR de ambas vueltas ($RIR_2 - RIR_1$), e informa sobre cómo ha variado el consenso, determinando si una ronda adicional propiciaría un mayor nivel de acuerdo entre expertos. Su ventaja es que trabaja con la mediana y los cuartiles. En la literatura se han encontrado distintos baremos observando que el criterio más seguido es $VRIR \leq .50$ (Arregui et al., 1996; I. Pérez et al., 2009; Soria y Martínez, 2010). A pesar de ello, se ha decidido utilizar en el presente estudio criterios más restrictivos, en la línea de los utilizados por Climent, Melián y Sanchis, (2014) y Seguí y Server (2010) donde la estabilidad se asume cuando $-.25 \leq VRIR \leq .25$.

b. Variación del Coeficiente de Variación (v).

Se obtiene de deducir el coeficiente de variación de la segunda ronda del coeficiente de la primera ronda ($CV_2 - CV_1$). No es aconsejable en distribuciones asimétricas (I.M. Martínez, 2003). Su principal inconveniente es que utiliza la media y la desviación típica, motivo por el que es aplicado como complemento del $VRIR$. Los baremos también son establecidos de modo arbitrario por los investigadores utilizando la lógica del sentido común de los resultados y el criterio estadístico (Ortega, 2008). De ese modo se hallan diversos valores de referencia en la literatura, por ejemplo, I. Pérez et al. (2009) y Von der Gracht (2012) establecen que si $v \leq .50$ el nivel de consenso es bueno y no se precisarían rondas adicionales. Siguiendo los mismos criterios que para $VRIR$, se opta por establecer criterios más restrictivos utilizando como linde de estabilidad el valor de .27 (Soria y Martínez, 2010).

4.3.4.2. Prueba de Wilcoxon.

La prueba de Wilcoxon (1945) es un análisis no paramétrico similar al T-test de Student. Pero a diferencia de éste, se utiliza cuando no se asume la normalidad en los datos, como es el caso. Se aplica para comprobar si la tendencia central (la mediana) de dos muestras difiere significativamente o no, mediante p-valor. Si $p\text{-valor} < 0.05$ se puede interpretar que las muestras que se comparan son diferentes con una probabilidad mayor del 95% (Gómez-Biedma, Vivó y Soria, 2001). En este estudio se aplicará para identificar el nivel de consistencia entre la opinión de los expertos en ambas rondas R1 y R2.

4.3.4.3. Coeficiente Kappa de Cohen

La prueba de concordancia de Kappa (k) refleja el grado de acuerdo entre los expertos, es muy utilizada para medir la concordancia entre dos observaciones pues tiene como principal ventaja que corrige el porcentaje de acuerdo debido al azar.

El rango de posibles valores de kappa es de -1 a 1, aunque por lo general se encuentra entre 0 y 1. La unidad representa un acuerdo perfecto, lo que indica que los evaluadores están de acuerdo en su clasificación de todos los casos. Cero indica acuerdo no es mayor que el esperado por el azar, como si el evaluador hubieran "adivinado" cada calificación. Una kappa negativa indicaría que el acuerdo es inferior al esperado por el azar (Sim y Wright, 2005, p. 259).

En 1977, Landis y Koch propusieron unos parámetros que han sido muy utilizados en ciencias sociales y, por ello, son los que se han seleccionado como referentes para establecer el grado de consonancia (ver tabla 34):

Tabla 34: Valores de Kappa
Fuente: Adaptado de Landis y Koch (1977)

Índice Kappa	Fuerza de concordancia
0.00	Pobre
0.01 - 0.20	Leve
0.21 - 0.40	Aceptable
0.41 - 0.60	Moderada
0.61 - 0.80	Considerable
0.81 - 1	Casi perfecta

4.4. RESULTADOS

4.4.1. Resultados 1ª ronda ítems cuantitativos

En este apartado se presentan los resultados de la primera iteración de cuestionarios, una vez analizados mediante la aplicación de los estadísticos descritos en el apartado anterior. En primer lugar, en la tabla 35 se muestran las distribuciones de frecuencias de las respuestas individuales, ordenadas por el grado de mayor acuerdo y expresadas como porcentaje del total de frecuencias. Las respuestas correspondientes a los ítems 10 y 17 tienen formato tanto cualitativo como cuantitativo por lo que cada pregunta se divide en dos apartados: "a" en formato de respuesta abierta y "b" que corresponde a la respuesta cerrada en escala tipo Likert.

Se ha eliminado del análisis cuantitativo de la primera ola el ítem 17b por la baja tasa de respuesta obtenida (n=5) y será objeto de análisis en la segunda ronda junto con el resto de ítems en formato abierto.

Tabla 35: Porcentaje de acuerdo entre panelistas R1 (N=16)

Fuente: Elaboración propia

ítem	A Favor [C4-C5]	En Contra [C1-C2]	ítem	A Favor [C4-C5]	En Contra [C1-C2]
% de acuerdo			% de acuerdo		
A1	100.00	0.00	A2	68.75	12.50
A4	100.00	0.00	A12	68.75	6.25
A6	93.75	0.00	A15	68.75	6.25
A7	93.75	0.00	A22	62.50	25.00
A3	87.50	0.00	A19	62.50	18.75
A8	87.50	0.00	A10b	56.25	31.25
A13	86.67	0.00	A11	56.25	31.25
A9	81.25	6.25	A21	50.00	18.75
A20	81.25	0.00	A23	50.00	6.25
A5	75.00	0.00	A14	37.50	6.25
A24	75.00	0.00	A25	37.50	25.00
A18	73.33	0.00	A16	31.25	43.75

El proceso seguido para analizar los datos de la primera ronda, tiene como primer paso hallar los ítems que no cumplen con los criterios de consenso preestablecidos (véase tabla 31). Una vez aplicados los correspondientes análisis estadísticos, se observa que 9 ítems (37.50%) no lograron alcanzar un acuerdo aceptable, mientras que 15 (62.50%) fueron consensuados (ver tablas 36 y 37).

Tabla 36: Resultados estadísticos de la primera ola: Criterios de no consenso

Fuente: Elaboración propia

Ítems	N	Q1	Q2 Me	Q3	RI	Q	RIR	σ	CV	A favor [C4-C5]	En contra [C1-C2]	Grado de acuerdo
Cuartiles										% acuerdo		
A10b	16	2.00	4.00	5.00	3.00	1.50	0.75	1.45	40.10	56.25	31.25	SIN CONSENSO
A11	16	2.00	4.00	5.00	3.00	1.50	0.75	1.45	40.13	56.25	31.25	
A14	16	3.00	3.00	4.00	1.00	0.50	0.33	0.87	26.36	37.50	6.25	
A16	16	2.00	3.00	4.00	2.00	1.00	0.67	1.00	33.97	31.25	43.75	
A19	16	3.00	4.00	4.00	1.00	0.50	0.25	1.15	32.36	62.50	18.75	
A21	16	3.00	3.50	4.00	1.00	0.50	0.29	0.89	26.22	50.00	18.75	
A22	16	2.25	4.00	4.00	1.75	0.80	0.44	1.15	33.54	62.50	25.00	
A23	16	3.00	3.50	4.00	1.00	0.50	0.29	0.81	22.85	50.00	6.25	
A25	16	2.25	3.00	4.00	1.75	0.88	0.58	1.08	32.55	37.50	25.00	

Me = Mediana / Q_x = Cuartiles / Q = Semi intercuartil / RI = Rango Intercuartílico / RIR = Rango Intercuartílico Relativo / σ = Desviación típica / CV = Coeficiente de Variación

La parte del cuestionario que ha acumulado la mayoría de discrepancias es la referida al apartado IV del cuestionario: “Corrección e interpretación de los resultados”, que engloba los ítems A19, A21, A22, A23 y A25. De los 9 reactivos sin consenso, resulta revelador que 6 de ellos (66,67%) hacen referencia a los sesgos de respuesta.

La siguiente fase trató de establecer el grado de consenso de los 15 ítems restantes que alcanzaron acuerdo, conforme a las normas de decisión que se describieron en párrafos precedentes (ver tabla 31). Como muestra la tabla 37, el tratamiento estadístico aplicado

muestra que 9 preguntas alcanzaron un amplio consenso de respuestas a favor, y 6 ítems con consenso suficiente a favor. No se halló ningún ítem con grado de acuerdo en contra. De hecho, la mediana para todos estos ítems fue ≥ 4 .

Tabla 37: Resultados estadísticos de la primera ola: Grado de consenso
Fuente: Elaboración propia

Ítems	N	Q1	Q2	Q3	RI	Q	RIR	σ	CV	A	favor	En contra	Grado de consenso a favor
			Me							[C4,C5]		[C1,C2]	
		Cuartiles								% acuerdo			
A1	16	4.25	5.00	5.00	0.75	0.38	0.15	0.45	9.42	100.00		0.00	Amplio
A2	16	3.00	4.00	4.75	1.75	0.88	0.44	0.98	25.73	68.75		12.50	Suficiente
A3	16	4.00	5.00	5.00	1.00	0.50	0.20	0.73	16.23	87.50		0.00	Amplio
A4	16	5.00	5.00	5.00	0.00	0.00	0.00	0.40	8.38	100.00		0.00	Amplio
A5	16	3.25	4.50	5.00	1.75	0.88	0.39	0.86	20.15	75.00		0.00	Suficiente
A6	16	4.00	5.00	5.00	1.00	0.50	0.20	0.63	14.05	93.75		0.00	Amplio
A7	16	4.00	5.00	5.00	1.00	0.50	0.20	0.63	14.05	93.75		0.00	Amplio
A8	16	4.00	4.50	5.00	1.00	0.50	0.22	0.72	16.43	87.50		0.00	Amplio
A9	16	4.00	4.50	5.00	1.00	0.50	0.22	0.93	21.90	81.25		6.25	Amplio
A12	16	3.00	4.00	4.75	1.75	0.88	0.44	0.89	22.84	68.75		6.25	Suficiente
A13	15	4.00	4.00	4.00	0.00	0.00	0.00	0.59	14.60	86.67		0.00	Amplio
A15	16	3.00	4.00	4.00	1.00	0.50	0.25	0.83	21.88	68.75		6.25	Suficiente
A18	15	3.00	4.00	4.00	1.00	0.50	0.25	0.64	16.55	73.33		0.00	Suficiente
A20	16	4.00	4.00	4.75	0.75	0.38	0.19	0.68	16.74	81.25		0.00	Amplio
A24	16	3.25	4.00	4.00	0.75	0.38	0.19	0.62	15.98	75.00		0.00	Suficiente

Me = Mediana / Q_x = Cuartiles / Q = Semi intercuartil / RI = Rango Intercuartílico / RIR = Rango Intercuartílico Relativo / σ = Desviación típica / CV = Coeficiente de Variación

El análisis estadístico de esta primera ola se empleó para ofrecer retroalimentación a los panelistas y poder confeccionar un segundo cuestionario destinado a promover el mayor consenso entre los expertos, donde éstos, una vez conocedores de los resultados, podían modificar su criterio si así lo consideraban. Las apreciaciones y conclusiones de los resultados serán descritas en el siguiente apartado.

4.4.2. Resultados 2ª ronda ítems cuantitativos

A continuación, se describe los resultados de la segunda circulación de cuestionarios, teniendo en cuenta que se prescinden de los datos obtenidos en las respuestas 10b y 17b por obtener baja tasa de respuesta $n \leq 4$. Estos ítems se analizarán en el epígrafe correspondiente al análisis de los ítems de formato abierto.

Al igual que se hizo en la primera ola, se presenta en formato de tabla la distribución de frecuencias de las respuestas individuales de cada experto en el segundo cuestionario, ordenadas por el grado de acuerdo (de mayor a menor valor) y expresada como porcentaje del total de frecuencias (ver tabla 38). Se observa que aumentó el grado de consenso a favor en cada uno de los ítems a excepción del ítem 20 (*“los sesgos son un problema importante pues podrían contaminar los resultados de medida de los test, pudiendo incluso invalidar toda la evaluación”*) que se mantiene invariable, y el ítem A13 (*“a la hora de aplicar un test suele darse más importancia a las propiedades psicométricas de fiabilidad y validez que a la detección de sesgos”*) donde se acrecentó la tasa de respuesta a favor de la indecisión.

Tabla 38: Porcentaje de acuerdo entre panelistas R2 (N=16)

Fuente: Elaboración propia

ÍTEMS	A Favor [C4,C5]	En Contra [C1,C2]	ÍTEMS	A Favor [C4,C5]	En Contra [C1,C2]
	% DE ACUERDO			% DE ACUERDO	
A1	100.00	0.00	A13	75.00	0.00
A3	100.00	0.00	A15	68.75	12.50
A4	100.00	0.00	A21	68.75	0.00
A6	100.00	0.00	A16	62.50	6.25
A7	100.00	0.00	A11	50.00	37.50
A8	93.75	0.00	A14	50.00	6.25
A2	87.50	0.00	A23	50.00	12.50
A5	87.50	0.00	A25	50.00	12.50
A19	87.50	6.25			
A9	81.25	6.25			
A12	81.25	0.00			
A18	81.25	0.00			
A20	81.25	0.00			
A22	81.25	0.00			
A24	81.25	6.25			

Siguiendo los mismos criterios y análisis estadísticos que en la R1, se identificaron en primer lugar los ítems que no superaron las reglas de decisión establecidas para el consenso en la tabla 31, y para el resto de ítems, se estableció el grado de consenso alcanzado según los baremos de la tabla 32. De ese modo, en esta 2ª ronda, la gran mayoría de panelistas volvió a reflexionar sobre todos los ítems, incluso los ya consensuados en la 1ª ola, minorando notablemente el porcentaje de respuestas dudosas o de indecisión (C3) y fortaleciendo sus posiciones al incrementar el grado de acuerdo a favor en todos y cada uno de los ítems ($Me \geq 4$), a excepción del ítem A20 que no experimentó ninguna variación y del ítem A13, donde aparecieron respuestas de acuerdo en contra.

De los 9 ítems sin acuerdo de la primera ola, los panelistas reconsideraron a favor 3; concretamente, los ítems A19, A21 y A22. Cabe recordar que la pregunta 10b se eliminó de este análisis para ser tratada posteriormente con el resto de datos cualitativos, por lo que finalmente quedaron 5 ítems sin lograr convergencia (A11, A14, A16, A23 y A25). Aunque sin alcanzar el mínimo de consenso establecido (67%), se observa que al menos el 50% de los expertos se postulan a favor (ver tabla 39). Para favorecer la comprensión, en la tabla 40 se muestra la redacción de los ítems no consensuados.

Tabla 39: Resultados estadísticos de la segunda ola: Criterios de no consenso

Elaboración propia

Ítems	N	Q1	Q2	Q3	RI	Q	RIR	σ	CV	A favor	En contra	Grado de acuerdo
		Cuartiles								[C4,C5]	[C1,C2]	
% acuerdo												
A11	16	2.00	3.50	5.00	3.00	1.50	0.86	1.67	0.49	50.00	37.50	SIN CONSENSO
A14	16	3.00	3.50	4.00	1.00	0.50	0.29	0.89	0.26	50.00	6.25	
A16	16	3.00	4.00	4.00	1.00	0.50	0.25	0.72	0.20	62.50	6.25	
A23	16	3.00	3.50	4.00	1.00	0.50	0.29	0.89	0.26	50.00	12.50	
A25	16	3.00	3.50	4.00	1.00	0.50	0.29	0.96	0.27	50.00	12.50	

Me = Mediana / Q_x = Cuartiles / Q = Semi intercuartil / RI = Rango Intercuartílico / RIR = Rango Intercuartílico Relativo / σ = Desviación típica / CV = Coeficiente de Variación

Tabla 40: Ítems sin consenso 2ª vuelta
Fuente: Elaboración propia

	ÍTEMS
A11	Es poco apropiado incluir en los cuestionarios de evaluación de riesgos psicosociales escalas cuya finalidad no es detectar riesgos, como son escalas de satisfacción o clima laboral
A14	La aquiescencia es uno de los sesgos que pueden atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales
A16	El autoengaño o la simulación son sesgos que pueden atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales
A23	A pesar de existir técnicas destinadas a detectar y minimizar los sesgos de respuesta a menudo son eludidos
A25	Es necesario tener conocimientos de psicometría para interpretar los resultados de las escalas de autoinforme utilizadas para valorar los riesgos psicosociales

Al igual que ha ocurrido en la primera ronda, los ítems A14, A16 y A23, referidos a los sesgos de respuesta, constituyen el punto crítico del disenso al sumar el mayor número de ítems sin acuerdo (83.33%); no obstante, se observa que más del 50% de expertos estuvieron a favor. Puede incluirse la cuestión A15 (sombreada en gris en la tabla 41), “*la deseabilidad social es uno de los sesgos que pueden atribuirse a escalas de auto-informe utilizadas en prevención de riesgos psicosociales*”), pues, aunque alcanza el consenso prefijado con una tasa de acuerdo a favor del 68.75% y un $RIR \leq 50$, sobrepasa tímidamente los límites de dispersión fijados ($\sigma = 1.02$, $CV = .28$). Sobre el ítem A25 sería de interés la detección de subgrupos, pues psicólogos y técnicos de prevención no cuentan con la misma formación en psicometría, y eso podría explicar la falta de consenso en este punto. Sin embargo, debido al reducido tamaño de la población muestral, el análisis resultaría infructuoso. En la tabla 41 se presentan los resultados de los ítems consensuados.

Tabla 41: Resultados estadísticos de la segunda ola: Grado de consenso
Fuente: Elaboración propia

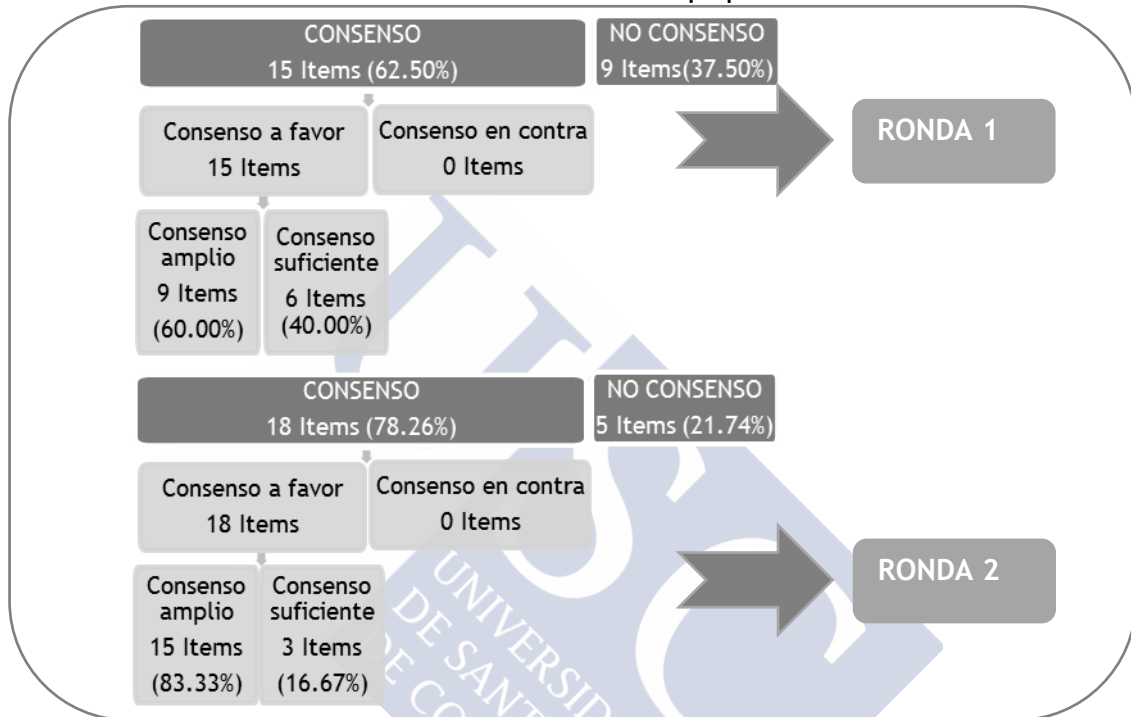
Ítems	N	Q1	Q2	Q3	RI	Q	RIR	σ	CV	A favor	En contra	Grado de consenso
		Me								[C4,C5]	[C1,C2]	
		Cuartiles								% acuerdo		
A1	16	5.00	5.00	5.00	0.00	0.00	0.00	0.40	0.08	100	0.00	Amplio
A2	16	4.00	4.00	5.00	1.00	0.50	0.25	0.68	0.16	87.50	0.00	Amplio
A3	16	5.00	5.00	5.00	0.00	0.00	0.00	0.40	0.08	100	0.00	Amplio
A4	16	5.00	5.00	5.00	0.00	0.00	0.00	0.25	0.05	100	0.00	Amplio
A5	16	4.00	4.00	5.00	1.00	0.50	0.25	0.70	0.16	87.50	0.00	Amplio
A6	16	5.00	5.00	5.00	0.00	0.00	0.00	0.40	0.08	100	0.00	Amplio
A7	16	4.00	5.00	5.00	1.00	0.50	0.20	0.50	0.11	100	0.00	Amplio
A8	16	4.00	5.00	5.00	1.00	0.50	0.20	0.63	0.14	93.75	0.00	Amplio
A9	16	4.00	5.00	5.00	1.00	0.50	0.20	0.96	0.22	81.25	6.25	Amplio
A12	16	4.00	4.00	4.75	0.75	0.38	0.19	0.68	0.17	81.25	0.00	Amplio
A13	16	3.25	4.00	4.00	0.75	0.38	0.19	0.68	0.17	75.00	3.25	Suficiente
A15	16	3.00	4.00	4.00	1.00	0.50	0.25	1.02	0.28	68.75	12.50	Suficiente
A18	16	4.00	4.00	4.75	0.75	0.38	0.19	0.68	0.17	81.25	4.00	Amplio
A19	16	4.00	4.00	4.75	0.75	0.38	0.19	0.77	0.19	87.50	4.00	Amplio
A20	16	4.00	4.00	4.75	0.75	0.38	0.19	0.68	0.17	81.25	4.00	Amplio
A21	16	3.00	4.00	4.00	1.00	0.50	0.25	0.58	0.15	68.75	3.00	Suficiente
A22	16	4.00	4.00	4.00	0.00	0.00	0.00	0.50	0.13	81.25	4.00	Amplio
A24	16	4.00	4.00	4.00	0.00	0.00	0.00	0.77	0.20	81.25	4.00	Amplio

Me = Mediana / Q_k = Cuartiles / Q = Semi intercuartil / RI = Rango Intercuartílico / RIR = Rango Intercuartílico Relativo / σ = Desviación típica / CV = Coeficiente de Variación

En la segunda vuelta, los porcentajes a favor se incrementaron siendo casi irrelevantes las posiciones en contra. Se observa además una notable disminución de los índices de dispersión, concediendo representatividad a la media aritmética y a la mediana. Así, cuando los juicios de los encuestados presentan un nivel aceptable de proximidad, se considera que existe consenso (Landeta, 2002). A efectos de comparación, en la figura 8 se establecen los porcentajes consenso y no consenso en cada una de las dos rondas.

Figura 8: Comparación entre rondas del resultado global

Fuente: Elaboración propia



Según los criterios establecidos, el consenso alcanzado en la última vuelta (R2) es de 18 ítems (casi el 80% del total), lo que constituye un grado de convergencia muy alto (Green, Jones, Hughes y Williams, 1999; Sumison, 1998). A su vez, 15 de estos 18 ítems alcanzaron el consenso de al menos 3/4 de los miembros, hecho considerado muy significativo.

Haciendo un resumen de las valoraciones de los 18 ítems que han sido objeto de convergencia de opinión, resalta la concordancia en afirmar que las evaluaciones de riesgos psicosociales son subestimadas por las empresas (A1, A2 y A4), y que existe confusión sobre lo que es un método de evaluación de riesgos (A3). Destaca también la importancia de utilizar herramientas que cumplan con los criterios de estandarización y baremación, además de las propiedades psicométricas de fiabilidad y validez (A6, A8 y A9). Los expertos también se ponen de acuerdo al valorar la escasez de los controles de sesgos y las herramientas para su control y la importancia de los sesgos de respuesta (A7, A12, A13, A15, A19, A2, A21, A22 y A24). También convergen al señalar que un protocolo de administración contribuiría a minimizar dichos sesgos (A18).

4.4.3.Resultados 2ª ronda ítems cualitativos

La riqueza de aportaciones realizadas por los panelistas en los ítems abiertos contribuyó a engrandecer la calidad del estudio, pues han sido una gran fuente de información basada en la experiencia y conocimientos de los expertos.

Las respuestas a las preguntas abiertas de la primera ronda se han presentado convertidas en 32 ítems cerrados en el segundo cuestionario permitiendo, de ese modo, que las reflexiones y argumentaciones realizadas a título individual por los panelistas, fuesen conocidas y valoradas por el resto de expertos, preservando en todo momento el anonimato de los miembros que contribuyeron con sus aportaciones adicionales. Para el análisis de su contenido se estructuraron a modo de inventario exhaustivo todas las respuestas recibidas, presentándolas con las propias palabras de los expertos para minimizar de este modo un posible efecto de interpretación. Los juicios con planteamientos similares se agruparon en un solo ítem. Seguidamente se clasificaron los ítems en categorías, atendiendo a la frecuencia de respuesta.

Se han dividido en dos bloques, por un lado se encuentran las respuestas a las cuestiones semicerradas A10a y A17a, que incluían la opción “otros” y, por otro lado, se hallan las aportaciones realizadas por los expertos en cada una de las dimensiones en las que se divide el cuestionario, incluyendo el apartado final de observaciones, comentarios y/o sugerencias.

Para cada ítem cualitativo, las variables se presentan ordenados primeramente por la media aritmética de forma decreciente (\bar{X}) y seguidamente por el menor rango intercuartílico relativo (RIR), se complementa la información con los porcentajes de acuerdo, tanto a favor como en contra y con la mediana (Me). Se considerarán las cuestiones como relevantes las que combinen valores de preferencia y consenso con alta significación, es decir, que cumplan los siguientes criterios (véase tabla 42).

Tabla 42: Criterios para considerar los acuerdos como significativos
Fuente: Elaboración propia

Orden	CRITERIOS PARA ALCANZAR UN ACUERDO SIGNIFICATIVO
1º	Me \neq 3
2º	Frecuencia de respuesta $\geq 75\%$
3º	RIR ≤ 0.5

Me = Mediana / RIR = Rango Intercuartílico Relativo

En todas las tablas referentes a los resultados de los ítems cualitativos, se mostrarán los resultados ordenados por la media aritmética, por ese motivo aparecerán sombreados en gris únicamente los ítems que cumplan con la totalidad de los criterios establecidos.

4.4.3.1.Bloque 1: Respuestas a las preguntas abiertas

El primer bloque de respuestas corresponde a las aportaciones realizadas por los expertos en la primera ronda a los ítems 10a y 17a. En el segundo cuestionario fueron presentadas a los expertos para su consiguiente valoración, cuyos resultados se muestran en las tablas 43 y 44.

Tabla 43: Resultados ítems cualitativos: Valoraciones ítem A10a
Elaboración propia

RESPUESTAS ABIERTAS AL ÍTEM 10a: Otros criterios no indicados en el cuestionario que debería contener la herramienta de recogida de información para ser lo más objetivos posible		Estadísticos			A favor - En contra [C4,C5] - [C1,C2]		N
		\bar{x}	RIR	Me	% acuerdo		
1º	Adaptado a la población a la que se dirige (tener en cuenta en empresas donde se encuentran trabajadores inmigrantes, discapacitados...)	4.47	0.25	4.00	1.00	0.00	15
2º	Tener en cuenta el entorno y al trabajador como miembro de una organización, no como entes aislados	4.20	0.25	4.00	0.93	0.00	15
3º	Ajuste al contexto de evaluación	4.15	0.13	4.00	0.92	0.00	13
4º	La participación y la autoconfrontación	4.15	0.25	4.00	0.85	0.00	13
5º	Las condiciones de administración	4.14	0.06	4.00	0.93	0.00	14
6º	(Pre)conocimiento del terreno o validez ecológica	4.00	0.13	4.00	0.79	0.00	14
7º	Debe tener medias de datos de otras empresas	3.07	1.00	3.00	0.29	0.21	14
8º	No utilización de escalas de Likert	2.29	0.25	2.00	0.07	0.57	14

\bar{x} = Media / Me = Mediana / RIR = Rango Intercuartílico Relativo

El ítem 10a “*Otros criterios no indicados en el cuestionario que debería contener la herramienta de recogida de información para ser lo más objetivos posible*”, es la cuestión con más aportaciones. De las ocho realizadas, seis de ellas han obtenido un alto porcentaje de acuerdo a favor. Los expertos se ponen de acuerdo para afirmar que hay que tener en cuenta al trabajador y su entorno laboral. La 8ª aportación “*no utilización de escalas Likert*” es la única de toda la batería de preguntas abiertas que ha alcanzado el mayor desacuerdo, un 57% en contra.

Tabla 44: Resultados ítems cualitativos: Valoraciones ítem A17a
Fuente: Elaboración propia

RESPUESTAS ABIERTAS AL ÍTEM 17 a: Otros sesgos de respuesta no indicados en el cuestionario que puedan atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales		Estadísticos			A favor - En contra [C4,C5] - [C1,C2]		N
		\bar{x}	RIR	Me	% acuerdo		
1º	En el día a día, muchas empresas siguen confundiendo conceptos tan diferentes como clima laboral o evaluación psicosocial	4.00	0.13	4.00	76.92	7.69	13
2º	Sesgo de baja tasa de respuesta (la muestra que se obtiene no es representa)	3.92	0.25	4.00	61.54	0.00	13
3º	La mala explicación conlleva sesgos de sobreestimar	3.85	0.50	4.00	76.92	7.69	13
4º	La utilidad estratégica de los datos para fines corporativos	3.67	0.38	4.00	58.33	8.33	12
5º	Excesiva distancia de lo medido	3.23	0.17	3.00	23.08	7.69	13

\bar{x} = Media / Me = Mediana / RIR = Rango Intercuartílico Relativo

El ítem 17a “*Otros sesgos de respuesta no indicados en el cuestionario que puedan atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales*” obtuvo 5 aportaciones, de las cuales sólo dos consiguieron el beneplácito de la mayoría de los expertos. La primera de ellas nos sitúa en una realidad compartida por la Inspección de Trabajo y Seguridad Social (2012) “*En el día a día, muchas empresas siguen confundiendo conceptos tan diferentes como clima laboral o evaluación psicosocial*”. El segundo aporte con mayor índice de consenso a favor se refiere al sesgo de sobreestimar, este sesgo surge cuando las condiciones de aplicación de los test no han sido apropiadas (Delgado, 1998), por lo que esta valoración

estaría mejor encuadrada en el apartado III de cuestiones abiertas, pues es el apartado específicamente dedicado a las cuestiones relativas a la aplicación de los test (ver tabla 47). Los tres ítems restantes, numerados en la tabla como 2º, 4º y 5º, no alcanzaron el 75% de respuestas en ninguno de los dos sentidos, ni a favor ni en contra.

4.4.3.2. Bloque 2: Listado de comentarios y sugerencias

En este bloque se hallan los aportes realizados por los expertos en el apartado de comentarios y sugerencias referidos a cada una de las cuatro dimensiones en las que se ha dividido el cuestionario, (i) Afirmaciones generales, (ii) Diseño de instrumentos de medida, (iii) Aplicación, y (iv) Corrección e interpretación de resultados. También se incluyen las observaciones reflejadas en el apartado final de comentarios y sugerencias (ver tablas 45 a 48).

Tabla 45: Valoraciones del apartado I: “Afirmaciones generales”
Fuente: Elaboración propia

RESPUESTAS ABIERTAS : Observaciones de la sección "Afirmaciones generales"		Estadísticos			A favor - En contra [C4,C5] - [C1,C2]		N
		\bar{x}	RIR	Me	% acuerdo		
1º	Hay una dejadez muy acusada en las empresas con respecto a las evaluaciones psicosociales porque las consideran subjetivas y poco útiles porque no se dan respuestas prácticas y rápidas a los problemas que se plantean	4.23	0.25	4.00	92.31	0.00	13
2º	El hecho de tener herramientas para valorar los riesgos psicosociales no quiere decir que los valoren más que otro tipo de riesgos. Pueden tener las herramientas pero no utilizarlas o hacerlo de forma incorrecta	4.00	0.25	4.00	76.92	0.00	13
3º	Lo importante de verdad es la metodología, más que el método	3.93	0.50	4.00	64.29	0.00	14
4º	En cuanto a la "2" y a la "3", depende del tipo y del tamaño de la empresa	3.75	0.25	4.00	66.67	8.33	12
5º	Encuentro diferencias entre las empresas que cuentan con evaluación de riesgos psicosociales y no: las de sectores como sanidad y socio-sanitario, principalmente, tienen implantados en mayor porcentaje, procedimientos para la gestión del riesgo psicosocial	3.71	0.25	4.00	64.29	7.14	14

\bar{x} = Media / Me = Mediana / RIR = Rango Inter cuartílico Relativo

De las cinco aportaciones efectuadas por los expertos en el apartado de comentarios de este bloque, sólo las dos primeras han cumplido con la totalidad de los parámetros establecidos para considerarlas como relevantes. Sin embargo, es destacable que el resto de comentarios obtuvo unos porcentajes de acuerdo a favor importantes, entre el 64.29% y el 66.67%.

Los ítems cualitativos correspondientes al apartado II del cuestionario “Diseño de instrumentos de medida los panelistas”, se presentan en la tabla 46.

Tabla 46: Valoraciones del apartado II: “Diseño de instrumentos de medida”
Fuente: Elaboración propia

RESPUESTAS ABIERTAS : Observaciones de la sección "Diseño de instrumentos de medida"		Estadísticos			A favor - En contra [C4,C5] - [C1,C2]		N
		\bar{x}	RIR	Me	% acuerdo		
1º	Un único instrumento de medida nos acerca al problema, pero para evaluar se requieren varias herramientas complementarias. No sólo la autopercepción del trabajador, sino del grupo, además de la organización del sistema y del trabajo. Debe haber como mínimo una triangulación de datos	4.67	0.20	5.00	100.00	0.00	15
2º	Algunos de los instrumentos han sido creados desde el entorno académico por lo que en muchos casos se han contentado con la obtención de datos y recomendaciones generales de actuación. Quienes venimos trabajando desde la intervención observamos que los instrumentos deben ser precisos y concretos respecto a las preguntas y respuestas. Ello facilita que en un retest puedan compararse los datos, observando y analizando las diferencias	4.15	0.13	4.00	92.31	0.00	13
3º	"Objetiva" conlleva un sesgo de objetivismo discutible. Acaso sería preferible "adecuada", "idónea"...A veces, los baremos dan sensación de "objetividad" y de "cientificidad", pero pueden conllevar efectos perversos de "naturalización" de condiciones de trabajo insanas. P. ej: puntuar por debajo de un estándar en "agotamiento" medido por MBI puede inducir a la gerencia de un centro a pensar que no estar tan mal como la media equivale a estar muy bien. Y a concluir que la "ciencia" dice que no hay nada que mejorar y que aún queda margen para sobrecargar al personal	3.69	0.25	4.00	61.54	7.69	13

\bar{x} = Media / Me = Mediana / RIR = Rango Intercuartílico Relativo

Los resultados de la tercera dimensión del cuestionario “Aplicación” se reflejan en la siguiente tabla (tabla 47). Los dos primeros comentarios son los únicos que, al cumplir con las tres medidas de inclusión fijadas de antemano, son considerados como relevantes. Resulta significativo que en la primera aportación un 100% de las respuestas lo hicieron para fijar el consenso a favor y en el segundo comentario, el consenso es establecido por el 92.31% de los panelistas que contestaron. Este apartado ha recibido cuatro aportaciones y su totalidad ha sido valorada por una amplia mayoría de los expertos, obteniendo un significativo consenso a favor, pues los porcentajes de respuesta se mueven entre el 85.71% y el 100%. Los dos primeros comentarios han sido objeto de unanimidad de respuesta, pues todos los expertos se posicionaron en la máxima puntuación a favor que corresponde al valor 5 (Totalmente de acuerdo), hecho que resulta trascendente.

Tabla 47: Valoraciones del apartado III: “Aplicación”
Fuente: Elaboración propia

Observaciones de la sección “Aplicación”		Estadísticos			A favor - En contra [C4,C5] - [C1,C2]		N
		\bar{x}	RIR	Me	% acuerdo		
1º	Es fundamental explicar el por qué, antes de comenzar la aplicación, además de cumplir con los estándares de aplicación de cualquier tipo de test	4.57	0.20	5.00	100.00	0.00	14
2º	Es importante también, la situación de la empresa en el momento en que se realiza la evaluación de riesgos (EREs, algún despido...), que la dirección esté visible durante alguna fase del proceso de la evaluación (presentación del estudio a los trabajadores, y solicitud de participación a todos...) y que quede bien garantizado el anonimato	4.54	0.20	5.00	100.00	0.00	13
3º	1) Es importante, previo a la administración, que el trabajador conozca la utilidad de la cumplimentación de la batería psicosocial y para que servirá. Es decir, que haya de antemano una buena información, y que la Dirección indique que tras el diagnóstico se intervendrá según los recursos y prioridades que se establezcan. 2) Influencia del contexto de la empresa / organización (presión agentes sociales, de la patronal, etc.)	4.53	0.20	5.00	93.33	0.00	15
4º	Conseguir sinceridad en las respuestas está en función, además de que esté bien hecha, de la forma de administración y de la credibilidad de la persona encargada de administrarlos	4.36	0.22	4.50	85.71	0.00	14

\bar{x} = Media / Me = Mediana / RIR = Rango Intercuartílico Relativo

La siguiente sección pertenece al IV bloque del cuestionario “Corrección e interpretación de los resultados” (ver tabla 48). Los comentarios realizados en este bloque no han adquirido acuerdos relevantes según los cánones fijados, pues ninguno alcanza la frecuencia 75%, no obstante, se debe tener en cuenta que, a excepción del último, han obtenido valoraciones de acuerdo a favor al menos un 50% de los participantes. Otro hecho a destacar en este apartado reside en que es el único en donde la mayoría de aportaciones refleja de manera notoria la diversidad de posiciones, pues existen valoraciones de acuerdo en contra que van desde el 15.38% al 38.46%.

Tabla 48: Valoraciones del apartado IV: "Corrección e interpretación de resultados"

Fuente: Elaboración propia

RESPUESTAS ABIERTAS : Observaciones de la sección: "Corrección e Interpretación de resultados"		Estadísticos			A favor - En contra [C4,C5] - [C1,C2]		N
		\bar{x}	RIR	Me	% acuerdo		
1º	Debe diferenciarse entre el control de sesgos de los instrumentos y el de la administración de la batería. Se debe ser consciente que en el campo de la prevención de riesgos, hay personas que trabajan en psicología aplicada a la prevención y no son psicólogos, por lo que se debe facilitar un nivel de interpretación para estas personas, mientras el contexto mundial en este campo no cambie, y que esta área pertenezca exclusivamente al campo de la psicología o afines.	3.77	0.38	4.00	69.23	15.38	13
2º	Los posibles sesgos producidos en el momento de administrar el cuestionario deben quedar recogidos por parte del administrador. Esto afectará a la interpretación de los resultados	3.73	0.25	4.00	66.67	0.00	15
3º	Se detectó en los técnicos de prevención de riesgos laborales (de una comunidad autónoma X) una insuficiente preparación para realizar con garantías de calidad, la corrección de cuestionarios de riesgos psicosociales, la interpretación de resultados y sus limitaciones, y la propuesta al empresario de medidas preventivas derivadas de tales evaluaciones.	3.40	0.71	3.50	50.00	20.00	10
4º	Inspección de trabajo prácticamente sólo te valida el método ISTAS como método válido para una evaluación psicosocial independientemente de que sea el más adecuado o no para la empresa en concreto.	2.92	0.67	3.00	30.77	38.46	13

 \bar{x} = Media / Me = Mediana / RIR = Rango Intercuartílico Relativo

La última parte de la encuesta es la referente al apartado de "Comentario final", se recibieron tres comentarios (tabla 49), siendo sólo uno valorado con acuerdo a favor por una amplia mayoría. El resto de aportaciones, aunque con frecuencias importantes del 50% y del 72.73%, no alcanzan el 75% establecido para ser consideradas como relevantes.

Tabla 49: Resultados ítems cualitativos: Valoraciones comentarios finales

Fuente: Elaboración propia

RESPUESTAS ABIERTAS : Observaciones de la sección: “Comentarios finales””		Estadísticos			A favor - En contra [C4,C5] - [C1,C2]		N
		\bar{x}	RIR	Me	% acuerdo		
1º	Fundamental e imprescindible, ampliar el proceso de evaluación de factores psicosociales, ir más allá de un cuestionario, y profundizar con los trabajadores en las condiciones y entorno en el que trabajan. Llevar a cabo sesiones de trabajo en las que los trabajadores puedan participar y hacerles partícipes del proceso. La temática psicosocial debería tener un mayor peso en las organizaciones, puesto que es el origen que no se ve, de otro tipo de problemáticas organizacionales.	4.38	0.25	4.00	92.31	0.00	13
2º	Como observación, decirte que para mí el problema, es que los cuestionarios de evaluación de los métodos más empleados son muy tediosos y muy poco adaptables a la realidad de las empresas y el trabajo del día a día...	3.82	0.25	4.00	72.73	9.09	11
3º	Pienso que para evitar centralidad en las respuestas, resulta más útil utilizar una escala de Likert con 6 posibles respuestas en vez de 5.	3.50	0.29	3.50	50.00	16.67	12

 \bar{x} = Media / Me = Mediana / RIR = Rango Intercuartílico Relativo

4.4.4.Resultados estabilidad individual

Mide la estabilidad de la opinión particular de cada experto en cada ronda. Se entendería como estabilidad absoluta que ninguno de los participantes modificase su respuesta en una determinada pregunta; hecho que no resulta deseable, pues lo esperado sería que la mayor parte de panelistas modifiquen su opinión una vez conocidos los resultados de la ronda anterior.

Han sido planteadas diversas medidas de cuantificación de este índice de estabilidad (Chaffin y Talley, 1980; Landeta, 2002), como el índice de asociación predictiva o el índice de expertos que modifican su valoración. Este último es uno de los más recomendados por resultar sencillo, de fácil empleo y aplicable a cualquier colectivo de expertos (Scheibe et al., 1975).

4.4.4.1.La proporción de expertos que modifican su valoración

Se entenderá adecuada la estabilidad cuando el porcentaje de expertos que modifiquen su respuesta sea menor a un nivel fijado previamente. Esta arbitrariedad en el establecimiento de baremos es su principal crítica. El indicador establecido como índice de referencia de estabilidad se ha fijado en 66% (Blanco, España, Franco y González, 2014); es decir, cuando las modificaciones de respuestas de los expertos no superen 1/3 de los participantes, puede interpretarse como satisfactorio. En esta investigación al no utilizar este criterio como finalización de iteraciones, los resultados no constituirán ningún criterio de decisión, sino que serán meramente informativos.

La fórmula de cálculo es la siguiente:

$$\text{Proporción Expertos que modifican su valoración} = (\text{N}^\circ \text{ de expertos que modifican su respuesta} / \text{N}^\circ \text{ total de expertos}) / 100$$

4.4.5.Resultados de estabilidad entre rondas

En el análisis de los resultados se consideró que se alcanzó un grado de estabilidad suficiente entre las rondas 1 y 2, cuando medida de diferentes formas, los valores fueron inferiores a los niveles previamente establecidos. Se eliminaron del tratamiento estadístico (ver tabla 50) los reactivos 10 y 17, al igual que se hizo para el establecimiento del consenso, por ser ambos ítems mixtos.

Comenzando en primer lugar por el análisis de la variación de los dos coeficientes de dispersión (VRIR y v), y teniendo en cuenta los parámetros fijados en la tabla 33, se observa que a excepción de los ítems A16, A22 y A25, que sobrepasan los índices preestablecidos, el resto de ítems (86.96%) reflejan índices de estabilidad muy reducidos. Los signos negativos señalan un incremento en la variación lo que, si bien ocurre en cinco de los ítems (21.74%) con valores casi inapreciables, revela que la técnica ha servido para hacer madurar las opiniones de los expertos participantes.

Tabla 50: Resultados de los medidores de estabilidad
Fuente: elaboración propia

Ítem	Variación coeficientes dispersión			Wilcoxon		Kappa	% variación respuesta Expertos	
	VRIR	Δ/∇	v	Δ/∇	(z)	(p)		(K)
A1	0.15	▽	0.01	▽	-1.00	0.32	0.82	6%
A2	0.19	▽	0.10	▽	-1.93	0.05 *	0.34	44%
A3	0.20	▽	0.08	▽	-2.24	0.02*	0.30	31%
A4	0.00	▽	0.03	▽	-1.00	0.32	-0.10	25%
A5	0.14	▽	0.04	▽	-0.38	0.71	0.61	25%
A6	0.20	▽	0.06	▽	-2.24	0.02*	0.34	31%
A7	0.00	▽	0.03	▽	-1.41	0.16	0.75	13%
A8	0.02	▽	0.03	▽	-1.73	0.08	0.67	19%
A9	0.02	▽	0.00	▽	-1.73	0.08	0.69	19%
A11	-0.11	△	-0.09	△	-2.00	0.05*	0.66	25%
A12	0.25	▽	0.06	▽	-1.73	0.08	0.71	19%
A13	-0.19	△	-0.03	△	-1.00	0.32	0.88	13%
A14	0.05	▽	0.00	▽	-0.54	0.59	0.38	38%
A15	0.00	▽	-0.06	△	-0.33	0.74	0.42	38%
A16	0.42	▽	0.14	▽	-2.59	0.01*	0.33	50%
A18	0.06	▽	0.00	▽	-1.41	0.16	0.41	38%
A19	0.06	▽	0.13	▽	-1.93	0.05*	0.34	44%
A20	0.00	▽	0.00	▽	0.00	1.00	0.79	13%
A21	0.04	▽	0.11	▽	-1.90	0.06	0.30	44%
A22	0.44	▽	0.21	▽	-1.73	0.08	0.26	44%
A23	0.00	▽	-0.03	△	-0.30	0.76	0.26	50%
A24	0.19	▽	-0.04	△	-0.58	0.56	0.66	19%
A25	0.30	▽	0.05	▽	-1.27	0.21	0.40	44%

*p-valor ≤ 0.05

VRIR = Variación rango intercuartílico relativo / v = variación coeficiente de variación

En las figuras 9 y 10 se puede observar gráficamente las variaciones de los panelistas para cada una de las dos rondas, donde se muestran los datos de los rangos intercuartílicos relativos y los coeficientes de variación.

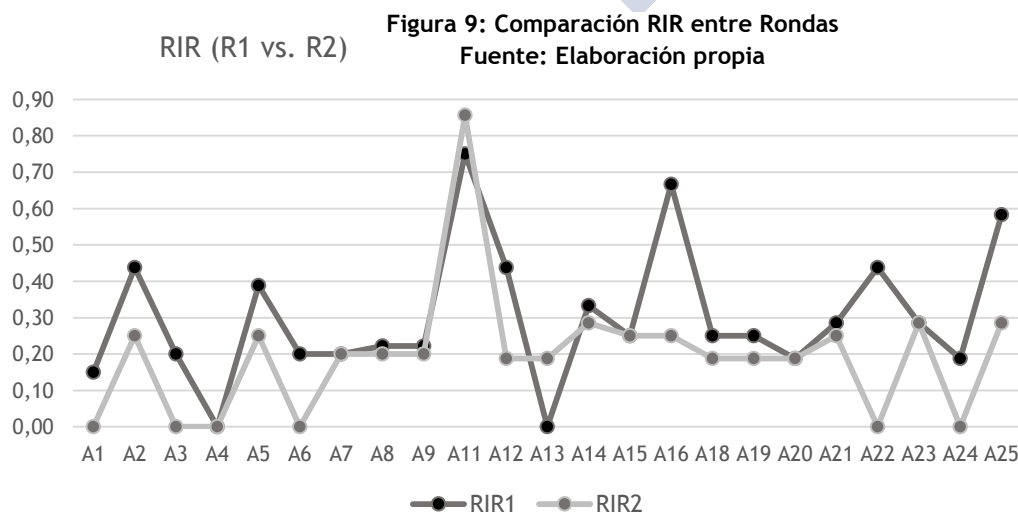
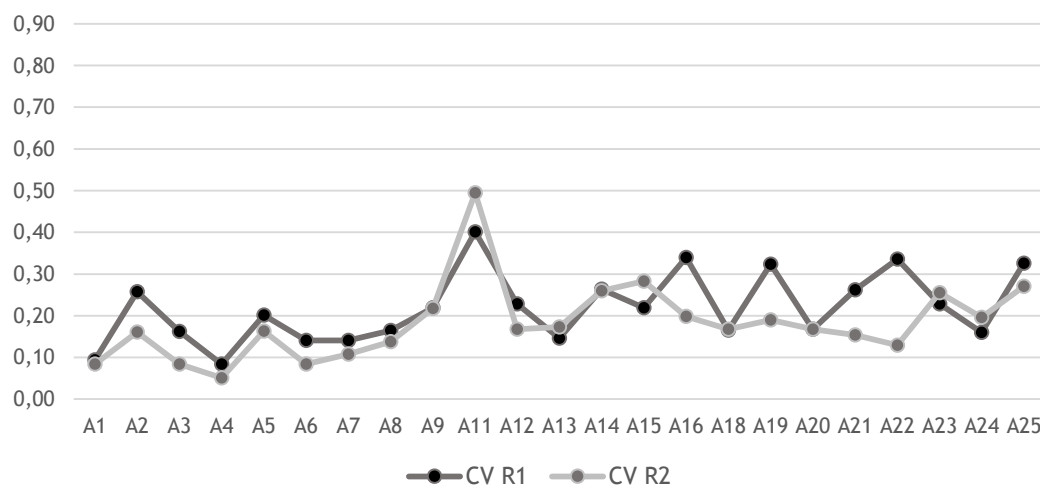


Figura 10 Comparación ν entre Rondas
 ν (R1 vs R2)
Fuente: Elaboración propia



De acuerdo con la prueba de Wilcoxon para muestras no simétricas de comparación de medianas, los resultados señalan sólo 5 cuestiones (21.74%) con diferencias significativas (A2, A3, A6, A16 y A19). Por su parte, siguiendo los indicadores de Landis y Koch (1977) para la prueba Kappa, solamente en el ítem A4 no se consigue una concordancia aceptable. Los reactivos A2, A3, A6, A14, A16, A19, A21, A22, A23 y A25 manifiestan concordancia aceptable, los ítems A15 y A18 revelan una concordancia moderada, los ítems A5, A7, A8, A9, A11, A12, A15 y A20 consiguieron una concordancia considerable y, finalmente, los ítems A1 y A13 alcanzaron un resultado de concordancia casi perfecto. Estos resultados demuestran que no existe una variación significativa entre rondas, más bien se evidencia una amplia estabilidad grupal.

En lo referente a la estabilidad individual, como era de esperar, todos los expertos modificaron en la segunda ronda, al menos su posición en uno de los ítems, pues se ofreció la posibilidad de variar su postura en todos y cada uno de los ítems del cuestionario, aunque ya estuviesen consensuados. Teniendo en cuenta que la mayoría de modificaciones tienen lugar en la segunda vuelta (Dodge y Clark, 1977), este índice tiene más sentido a partir de la tercera ronda, donde lo esperado es que sean menos los participantes que varíen sus posiciones.

4.5. DISCUSIÓN DE LOS RESULTADOS Y CONCLUSIONES

El presente estudio pretendía esbozar una composición de lugar de la problemática que envuelve las evaluaciones de riesgos psicosociales en el trabajo y, al mismo tiempo, explorar las principales debilidades y fortalezas del uso de cuestionarios en la medición de dichos riesgos. Para ello, a través de un estudio Delphi, fueron consultados 16 reputados expertos del ámbito académico y profesional. Se apostó por prefiar dos rondas de antemano (Landeta, 2002), cuestión ésta que resultó acertada, dada la notable estabilidad grupal que reflejan los resultados; en otras palabras, si se hubiera utilizado este criterio como finalización no habría sido necesario una tercera vuelta, reforzando así la decisión de haber optado por sólo dos iteraciones. La calidad técnica del panel de expertos, tanto respecto al tamaño como a los criterios de selección empleados, la inexistencia de casos de abandono, el tiempo invertido, la

producción por parte de los panelistas de ítems cualitativos, y el respeto del anonimato, entre otros, avalan el rigor del estudio que ha respetado las directrices marcadas por los padres del método Delphi, Norman Dalkey y Olaf Hermes (1963), y por otros investigadores como Theodore J. Gordon (1994).

En líneas generales, los resultados señalan un elevado grado de coincidencia en las opiniones de los expertos en la mayoría de las cuestiones formuladas. Además, en conjunto, las valoraciones de los participantes presentan una notable estabilidad y, dado que tanto los reactivos cuantitativos como los cualitativos superan el 60% de convergencia, se puede sugerir que académicos y profesionales comparten opinión acerca de los entresijos de la Ergonomía y Psicología Aplicada. Para detallar los hallazgos, se dividirá la discusión de los resultados en 2 bloques. Por un lado, se abordarán los ítems que alcanzaron un grado de acuerdo significativo ordenados según la estructuración del cuestionario, y por otro lado se discutirán los puntos críticos en donde los expertos no consiguieron ponerse de acuerdo, ni a favor ni en contra.

4.5.1. Ítems consensuados

Como se ha anticipado, los expertos consultados han mostrado un notable grado de consenso. De hecho, cerca del 80% de los ítems presentados han alcanzado un consenso favorable. Incluso, 15 ítems alcanzaron convergencia por una amplia mayoría; es decir, al menos 3/4 de los panelistas se mostraban “de acuerdo” o “totalmente de acuerdo” con el enunciado presentado, lo que hace que los resultados sean valorados como muy positivos.

4.5.1.1. Apartado de afirmaciones generales (Ítems A1-A4)

Con relación al contenido de estos ítems, cabe destacar que los expertos han coincidido con los principales postulados de la literatura revisada, al señalar por unanimidad que las evaluaciones de riesgos psicosociales son subestimadas por las empresas (ítems A1), en línea con (1) los argumentos de otros investigadores como Barreiro (2006), Cifre et al. (2011), M. García y Olarte (2010), Leka y Jain (2010), Merino (2013) o Cifre et al. (2009), que catalogan la evaluación de los riesgos psicosociales como *la asignatura pendiente*; y (2) con las principales estadísticas oficiales europeas y nacionales (ENCT VII, 2011; ESENER, 2010; INSHT, 2009) donde sale a luz, por ejemplo, el pequeño porcentaje de empresas europeas que intentan reducir los niveles de estrés (26%), o la diferencia de recursos y esfuerzos que se dedican para reducir los riesgos físicos o mecánicos en detrimento de los psicosociales, sirva de ejemplo el 51.70% dedicado a estudios sobre movimiento repetitivos o esfuerzos físicos, frente el 28.80% dedicado a pausas, horarios o estrés (ENCT VII, 2011).

Dotando de más coherencia al resultado de consenso en este ítem, se encuentra la aportación realizada (en el apartado I de Afirmaciones Generales del cuestionario) por uno de los panelistas, quien afirma que existe *“una dejadez muy acusada en las empresas con respecto a las evaluaciones psicosociales porque las consideran subjetivas y poco útiles porque no se dan respuestas prácticas y rápidas a los problemas que se plantean”*. Esta valoración fue apoyada por más de 92% de los expertos y estaría reforzando los hallazgos de Igartúa et al. (2006, citado por I. Cortés et al, 2010) quienes, a través de un Delphi realizado a médicos del

trabajo, éstos señalaron que uno de los principales obstáculos de la gestión de los riesgos psicosociales es debida a la reticencia de los directivos.

Este es un resultado inquietante, ya que desde el punto de vista legal existe la misma obligación por parte de las empresas de evaluar los riesgos psicosociales que, por ejemplo, los riesgos higiénicos y mecánicos tradicionales (Meliá, 2008). Entonces, cabe preguntarse qué es lo que falla. Una posible explicación reside en que las empresas cuentan con escasas herramientas para gestionar los riesgos psicosociales (ítem A2). De hecho, de acuerdo con la encuesta ESENER-2 (2015), tan sólo 1/3 de los centros de trabajo con más de veinte trabajadores en España afirma contar con algún procedimiento específico:

La encuesta demuestra que los riesgos psicosociales se consideran un reto mayor y más difícil de gestionar que los riesgos «tradicionales» para la seguridad y la salud en el trabajo. Es preciso sensibilizar a la gente y disponer de herramientas prácticas y sencillas que faciliten la gestión del estrés, la violencia y el acoso relacionados con el trabajo (EU-OSHA 2016, sección de ¿Es importante el problema?, párr.2).

Hay que tener en consideración que estamos tratando con riesgos de muy reciente incursión en el mundo de la prevención si se compara con los riesgos de seguridad o con los riesgos higiénicos. En España, el camino se está recorriendo espaciosamente, pues es uno de los pocos países de la UE que no tiene implantados criterios legales sobre la forma de evaluarlos (M. García y Olarte, 2010; Cifre et al., 2009). Estas circunstancias han dado cabida a un amplio abanico de herramientas de evaluación si bien, desde un punto de vista estrictamente psicométrico, algunas pueden ser cuestionadas como se desprende de la revisión efectuada dentro del marco teórico. Además, tal y como reafirma un 77% de los panelistas, *“el hecho de tener herramientas para valorar los riesgos psicosociales no quiere decir que los valoren más que otro tipo de riesgos. Pueden tener las herramientas, pero no utilizarlas o hacerlo de forma incorrecta”*.

Pero los riesgos psicosociales no sólo están subestimados, sino que también están envueltos en cierta confusión terminológica (R. Fernández, 2010, p. 167; Oncins y Almodóvar, 1997). Por ejemplo, entre lo que es un método y lo que es una técnica de evaluación de riesgos, pues a menudo se piensa que la evaluación de riesgos psicosociales consiste en aplicar un simple cuestionario (ítems A3 y A4). Como señala Niño (2006, p. 34), *“la utilización de las encuestas y las entrevistas a los trabajadores es el principal «hilo conductor» de las comprobaciones a realizar sobre las condiciones de trabajo, pero no constituye el fin en sí mismo de las evaluaciones psicosociales”*. Una posible solución se encuentra en Guàrdia y Però (2010) quienes, para dejar de alimentar el entuerto, recomiendan evitar referirse a métodos de evaluación de riesgos psicosociales cuando lo que realmente se hace es debatir sobre el rigor científico de las técnicas y los procedimientos propuestos para la recogida de datos.

4.5.1.2. Apartado de diseño de instrumentos de medida (Ítems A5-A11)

Una gran mayoría de participantes del panel (87.50%) ha manifestado lo importante que es la autopercepción del trabajador para poder valorar los riesgos psicosociales (ítem A5). Sin las observaciones realizadas desde el punto de vista del propio trabajador, que es quien experimenta

los riesgos, no se podría tener acceso a lo que Barker et al. (2002) denominan “información fenomenológica”; es decir, “la percepción de los encuestados de sí mismos y de su mundo, imposibles de conseguir de ningún otro modo” (p. 95). Dada la invisibilidad de muchos de los estresores laborales, sería muy difícil de realizar una evaluación de riesgos psicosociales sin contar con la percepción del trabajador sobre el grado en que está expuesto a factores de riesgo o sobre su propia experiencia de estrés (Moreno y Báez, 2010; Peiró, 2008; Villalobos, 2004), ya que al fin y al cabo es quien lo sufre. Pero las evaluaciones de riesgos psicosociales deben trascender a las percepciones individuales. Como han indicado el 100% de los panelistas no sólo “*la autopercepción del trabajador, sino del grupo, además de la organización del sistema y del trabajo. Debe haber como mínimo una triangulación de datos*”. Precisamente, la triangulación metodológica es una de las principales recomendaciones para obtener evaluaciones precisas y certeras. Según J. Morse (1991) consiste en la combinación de dos o más metodologías, generalmente cualitativa y cuantitativa, para obtener una aproximación más real del fenómeno objeto de estudio. Es una de las medidas más apreciadas en la literatura para dotar de mayor calidad a la evaluación de riesgos; concretamente se recomienda utilizar las entrevistas como complemento de los cuestionarios de autoinforme (Artacoz y Molinero, 2004; EU-OSHA, 2005; Ferrer, 2010; Guàrdia y Però, 2010; ITSS, 2012; Meliá et al., 2006; Moreno y Báez, 2019, 2010; C. Nogareda y Almodóvar, 2006; Oncins y Almodóvar, 1997).

Relacionados con la autopercepción se hallan los autoinformes, que son las herramientas más utilizadas para evaluar los riesgos psicosociales (ITSS, 2012; Moorman y Podsakoff, 1992) y que a su vez también son criticadas por una parte de la comunidad científica, que cuestiona precisamente la subjetividad de sus resultados (Fernández-Ballesteros, 2013). Para salvar la falta de objetividad y poder cuantificar los datos, se exige que estas herramientas cumplan con una serie de criterios (ITSS, 2012; RSP, 1997) como, por ejemplo, una adecuada fundamentación teórica, propiedades psicométricas de fiabilidad y validez; control de sesgos o baremación y estandarización. En este sentido, los resultados del Delphi fueron contundentes: cumplir con las propiedades psicométricas de fiabilidad y validez (ítem 6) y controlar los sesgos (ítem 7) obtuvieron el 100% de acuerdo a favor; la necesidad de disponer de baremación y estandarización (ítem 8) el 93.75%; y, por último, una adecuada fundamentación teórica (ítem 9) el 81.25%. Con todo, estos resultados no suponen ninguna revelación importante, pues están en consonancia con la lógica psicométrica ya que de ellos depende la calidad de la evaluación de riesgos (Abad, et al., 2011; Guàrdia, 2010a, 2010b; Meliá, 2008; Meliá et al., 2006; Peiró, 2010). A su vez, coinciden desde el punto de vista legal con los fundamentos de la *Guía de Actuaciones de la Inspección de Trabajo y Seguridad Social sobre Riesgos Psicosociales* (ITSS, 2012) y con el Art. 5.2 del RSP, que establecen que el procedimiento de evaluación deberá proporcionar confianza sobre su resultado y, en el caso de utilizar técnicas cuantitativas como los cuestionarios de autoinforme, se tendrá que valorar: (i) la fundamentación teórica, (ii) la calidad psicométrica del instrumento (fiabilidad y validez), (iii) el análisis de los ítems, y (iv) la baremación.

Indudablemente, estas cuestiones no agotan los criterios específicos para salvaguardar la calidad de las evaluaciones de riesgos psicosociales. De hecho, al ser preguntados por “*Otros criterios no indicados en el cuestionario que debería contener la herramienta de recogida de información para ser lo más objetivos posible*” (ítem A10) los expertos manifiestan de forma

unánime la importancia de adaptar el cuestionario a la población a la que se dirige, por ejemplo, teniendo en cuenta si existen en la empresa trabajadores inmigrantes o discapacitados. En este sentido los test tienen que ser justos y no “beneficiar a ciertos grupos de población en detrimento de otros de igual nivel en el rasgo que interesa medir” (Gómez- Benito et al, 2010, p. 76). Por ejemplo, un trabajador procedente de otro país puede malinterpretar los enunciados del cuestionario al atribuir un significado equivocado a las palabras. En la misma línea, mayoritariamente apuntaron la necesidad de ajustar el cuestionario al contexto de evaluación. Esto se refiere a la conveniencia de no utilizar escalas genéricas para toda la organización, pues cada puesto contiene una realidad psicosocial diferente, y sería preciso evaluar por puestos de trabajo o por colectivos de trabajadores (C. Nogareda y Almodóvar, 2006).

También señalan que deben cuidarse las condiciones de administración de los cuestionarios pues de ellas depende en gran medida la calidad de los resultados; desde un local con las condiciones térmicas inadecuadas, el ruido, la ansiedad de los trabajadores al ser encuestados, una preparación inadecuada del administrador de los test y un largo etcétera de factores que contribuyen a distorsionar los resultados obtenidos y por ende las generalizaciones que se realicen a partir de ellos.

La participación en las actividades de prevención, incluida la evaluación, es otro aspecto que destacan como crucial. Además, es una obligación establecida por Ley (arts.18.2, 34 y 36.2, LPRL), y que se recoge de modo muy general tanto en la *Guía de Actuaciones de la Inspección de Trabajo y Seguridad Social sobre Riesgos Psicosociales* (ITSS, 2012), como en la NTP 450 (Oncins y Almodóvar, 1997).

Es necesaria también la participación de los trabajadores o sus representantes en todas estas etapas o fases de la evaluación. Los trabajadores o sus representantes deben ser consultados y deben participar desde la primera fase de la planificación respecto a qué cosas hay que evaluar, dónde y cómo, hasta la fase de decisión de las medidas que se deben adoptar para la mejora de las condiciones de trabajo y el control y seguimiento de dichas mejoras preventivas (p. 2).

Sin embargo no existen directrices en forma de manuales o guías de recomendación sobre cuál es el mejor modo de participación, y menos de cómo abordar la confrontación de los resultados, lo que puede contribuir a evaluaciones menos precisas.

Finalmente, también han aflorado cuestiones de debate que surgen sobre la base de la distinta procedencia profesional de los expertos consultados. En concreto, los profesionales reprochan que, con frecuencia, las herramientas de prevención creadas desde el entorno académico pecan de formular ítems y modos de actuación demasiado generales. Una posible solución para conseguir instrumentos más precisos en cuanto a preguntas y respuestas, tal y como reclaman desde la rama intervencionista, sería la cooperación y colaboración de todos los sectores implicados y preocupados por la prevención. De este modo, a la hora de confeccionar un método de evaluación se podría contar con la experiencia y puntos de vista de todos los profesionales implicados.

4.5.1.3. Apartado de aplicación (ítems A12-A18)

Los expertos están de acuerdo en que las medidas de autoinforme son vulnerables a la falseación (ítem A12). Cuestionan la validez de estas herramientas debido a su falta de objetividad y a lo sensibles que son a sesgos como la aquiescencia, la deseabilidad social o la disimulación, entre otros (Fernández-Ballesteros, 2013; González y Iruarrizaga, 2005; González-Llaneza, 2007; Jensen, 1995; y Navas, 1999). De hecho, De las Cuevas y González de Rivera (1992) son contundentes: *“los inventarios auto-administrados están especialmente sujetos a la simulación o falsificación por parte de quien los cumplimenta”* (p. 3). Cobra por ello una importancia fundamental cumplir con los criterios que garanticen la mayor objetividad de los instrumentos de recolección de datos, como por ejemplo la fiabilidad, la validez o el control de fuentes de error. Precisamente, los ítems relacionados con los sesgos han sido las cuestiones que ha presentado más controversia entre los expertos (ítems A14, A16 y A23). No obstante, para seguir un orden lógico, los resultados controvertidos serán abordados en último lugar.

Volviendo sobre los requisitos que todo instrumento de evaluación debe reunir para ser válido, se ha seguido preguntando a los participantes de panel, aunque ésta vez se ha querido conocer cuál es su experiencia acerca de cuáles son los criterios prioritarios a la hora de seleccionar un cuestionario válido. En ese sentido, un 75% reconoció que se concede más importancia a las propiedades psicométricas de los test que al control de sus sesgos (ítem A13). Ello va en línea con las afirmaciones de Dupuis et al. (2015) y con la revisión de cuestionarios de evaluación realizada en esta tesis, donde se recuerda que de todas las medidas exploradas sólo dos de ellas hacen mención explícita al control de sesgos (BFSL y MC Mutua). También es necesario destacar la cantidad de veces en las que, al hablar de propiedades psicométricas, el foco de atención gira sobre la necesidad de garantizar la fiabilidad y la validez, sin aludir a la importancia de controlar debidamente los sesgos (CEOE, 2013; Guàrdia, 2008b; Moreno y Báez, 2010).

Curiosamente, aunque los expertos han coincidido al afirmar que los autoinformes son susceptibles de ser falseados (ítem A12) y que el control de sesgos dotaría de mayor validez a los cuestionarios (ítem A7), cuando se les pregunta cuál de los sesgos concretos que pueden influir en las evaluaciones de riesgos (i.e., aquiescencia [A14], DS [A15], autoengaño o simulación [A16]), surgen las discrepancias y se desvanece el consenso. Como se trata de preguntas cerradas, los expertos no pudieron realizar en ese momento aclaraciones ni mencionar otro tipo de sesgos no referenciado en los ítems.

Salvando esa limitación, se encuentra el ítem 17 como pregunta abierta: *“Otros sesgos de respuesta no indicados en el cuestionario que puedan atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales”*. En este ítem los expertos han podido reflejar sin cortapisas las fuentes de error que, a su juicio, están incidiendo en la calidad de los cuestionarios de evaluación de riesgos psicosociales, siendo dos los aportes que han recibido el asentimiento de la mayoría de panelistas (76.92%). El primero de ellos asegura que muchas empresas aún no tienen claro en qué consiste una evaluación de riesgos psicosociales, pues a menudo confunden clima laboral con dicha evaluación. Este punto de vista es muy interesante, pues está reflejando la existencia de carencias informativas y formativas básicas en las organizaciones sobre el contenido de las evaluaciones de riesgos psicosociales. Esta cuestión

no ha pasado desapercibida para la ITSS (2012), quien de manera tajante subrayó que las escalas que no detecten riesgos, como las de clima laboral o satisfacción, no se pueden utilizar para realizar esa labor preventiva; en su caso, estarían bien utilizadas como herramientas complementarias, pero no como herramientas de evaluación de riesgos psicosocial, puesto que no lo son. La segunda de las aportaciones recibió el mismo nivel de consenso a favor, y trata sobre el sesgo de sobreestimar, directamente relacionado con las condiciones de aplicación inapropiadas (Delgado, 1998), que viene a reforzar la importancia de contar con un protocolo de aplicación de cuestionarios.

En este sentido, los expertos confirman la utilidad de contar con un protocolo de administración de cuestionarios de cara a minimizar los posibles sesgos de respuesta. Justamente de este protocolo se hace eco AERA, et al. (2014) e ITC (2010) como parte de los estándares de buenas prácticas relacionadas con el uso de los test en investigación. Como ejemplo de los puntos a incluir dentro de ese protocolo se encuentran (1) la formación y entrenamiento previo de la persona encargada de administrar los test; (2) la preparación con anterioridad del lugar dedicado a cubrir los cuestionarios, eliminando las posibles distracciones e incomodidades; (3) el fomento de una relación empática con los encuestados para intentar rebajar los niveles de ansiedad de las personas que van a cubrir los test; o (4) la lectura de las instrucciones de forma clara y contestando las posibles dudas que puedan surgir. El INSHT (2015) es otra de las organizaciones que recomienda la planificación de actuaciones antes de comenzar la evaluación de los riesgos psicosociales.

Asimismo, el protocolo de aplicación debería (1) cumplir con los estándares de aplicación de los test y explicar la utilidad de los cuestionarios antes de repartirlo, (2) implicar a la dirección durante al menos alguna fase del proceso de evaluación, y (3) tener en cuenta la situación de la organización en el momento de evaluar; por ejemplo, si la empresa está incurso en algún ERE o se esperan despidos los trabajadores estarán menos receptivos, y por lo tanto sus respuestas pueden llegar a estar sesgadas.

4.5.1.4. Apartado de corrección e interpretación de los resultados (Ítems A19-A25)

La mayoría de reactivos de este apartado del cuestionario guarda una lógica y estrecha relación con los ítems del epígrafe anterior. De este modo, si la mayoría de expertos indica que se da más importancia a la validez y a la fiabilidad de las medidas, en menoscabo del control de sesgos (ítem A13), es coherente que ratifiquen en un alto porcentaje (superior al 81%) que (1) no es frecuente realizar controles de sesgos cuando se evalúan riesgos psicosociales (ítem A19); y (2) que las herramientas de evaluación existentes apenas incluyen mecanismos para gestionar estas fuentes de error (ítem A22).

A lo largo del marco teórico se ha podido constatar la dimensionalidad del problema inherente a los sesgos, y es que pueden llegar a invalidar todas o gran parte de las conclusiones de una investigación (Cheung y Chan, 2002; Cheung y Rensvold, 1999; Choi et al., 2010; De la Guardia et al., 2012; Dupuis et al., 2015; Harzing et al., 2012; Moya, 2009; Muñiz, 1998b; Restrepo y Gómez-Restrepo, 2004; Tomás et al., 2012; Salamanca, 2013). Como era de esperar, los participantes del panel no son desconocedores de esta realidad, y en esa dirección se encaminó su juicio cuando el 81.25% categorizaron los sesgos como amenazas a la validez de los resultados de las evaluaciones de riesgos psicosociales (ítem A20). Además, estuvieron de

acuerdo en que los sesgos de respuesta son los de mayor incidencia en los cuestionarios de autoinforme, en consonancia con la comunidad científica (Adams, 1961; Barnette, 2000; Barker et al., 2002; Batchelor et al., 2013; Chami-Castaldi et al., 2008; Cosentino y Castro, 2008; Cronbach, 1946, 1950; Donaldson y Grant-Vallone, 2002; Dupuis et al., 2015; Echeburúa et al., 2003; Hui y Triandis, 1985, 1989; Jensen, 1995; F. Jiménez y Sánchez, 2004; R.E. McGrath et al., 2010; Morales, 2006; Nardi, 2003; Nunnally y Bernstein, 1995; Paulhus, 1981, 1984, 1986, 1991; Paulhus y Vazire, 2007; Viswanathan, 2005).

Aunque apenas existe literatura acerca de la percepción del trabajador sobre la vulneración del contrato psicológico, un 81% de expertos afirman que podría influir en el falseamiento intencional de las respuestas de las escalas (ítem A24), tal vez con objeto de causar daño a la empresa y contrarrestar así el perjuicio padecido por las expectativas incumplidas (Bies y Tripp, 1995; Robinson y Rousseau, 1994). Según varios autores está muy relacionado con la satisfacción laboral, la confianza en la organización y con la permanencia en la empresa, entre otras variables (Guest, 2004, Guest y Conway, 2002; Saldarriaga, 2013; Solanes y Zaragoza, 2010; Tena, 2002). Precisamente, por la escasez de estudios al respecto, sería interesante explorar esta dimensión como posible distorsionadora en las respuestas de los autoinformes.

4.5.1.5. Apartado de sugerencias y observaciones

De acuerdo con los expertos consultados, es *“fundamental e imprescindible, ampliar el proceso de evaluación de factores psicosociales, ir más allá de un cuestionario, y profundizar con los trabajadores en las condiciones y entorno en el que trabajan. Llevar a cabo sesiones de trabajo en las que los trabajadores puedan participar y hacerles partícipes del proceso. La temática psicosocial debería tener un mayor peso en las organizaciones, puesto que es el origen que no se ve, de otro tipo de problemáticas organizacionales”*. Dicha cuestión bebe de aportaciones realizadas en anteriores apartados del cuestionario, reafirmando así el poco peso con el que gozan las evaluaciones de riesgo psicosocial en las organizaciones, recomendando prestar atención tanto al trabajador como a su entorno laboral, así como contar con su participación y asesoramiento en el proceso de evaluación.

4.5.2. Ítems sin consenso

Una valiosa e importante fuente de información es la concerniente a los ítems discordantes. En primer lugar, se encuentra la pregunta A15 que se refiere a la deseabilidad social como sesgo potencialmente distorsionador en los cuestionarios de evaluación de riesgos psicosociales. Aunque esta cuestión sí sobrepasó el acuerdo mínimo (68,75%), las respuestas de los expertos tuvieron un alto índice de dispersión y por ese motivo no fue incluida con el resto de ítems consensuados. Por otra parte se encuentran los reactivos que no lograron un mínimo de acuerdo, ni a favor ni en contra, es decir, donde la respuesta mayoritaria fue la indecisión (ítems A11, A14, A16, A23 y A25).

Resulta paradójico que la mayor parte de estos ítems objeto de discrepancia, tanto en la R1 como en la R2, hagan referencia explícita a los sesgos de respuesta, concretamente a los sesgos de deseabilidad social (ítem A15), aquiescencia (ítem A14), autoengaño o simulación (ítem A16), y al propio control de los sesgos (ítem A23). Estos resultados tal vez sean debidos a la falta de estudios empíricos sobre los sesgos de respuesta en las evaluaciones de riesgos

psicosociales, hecho que puede haber favorecido la divergencia de posicionamientos. La heterogeneidad de los participantes en el estudio (académicos y profesionales), con su diferente formación sobre psicometría, puede haber influido en no alcanzar acuerdo (ni a favor ni en contra) en los puntos mencionados. Se supone que los académicos, como creadores de buena parte de las herramientas de evaluación, poseen conocimientos más profundos sobre construcción de test y por consiguiente de las amenazas a su validez, conocimientos que, por ejemplo, no se contemplan dentro de los programas formativos de los técnicos de prevención de riesgos psicosociales. De ser así, se entendería un diferente punto de vista acerca de la importancia de los sesgos de respuesta.

Los resultados del ítem A11 resultan curiosos. En este punto no ha habido convergencia de opinión sobre si es “apropiado incluir en los cuestionarios de evaluación de riesgos psicosociales escalas cuya finalidad no es detectar riesgos, como son escalas de satisfacción o clima laboral”; sin embargo, sí apoyaron por amplia mayoría una cuestión similar propuesta por uno de los panelistas: “*en el día a día, muchas empresas siguen confundiendo conceptos tan diferentes como clima laboral o evaluación psicosocial*” (ítem A17). En el intento de entender esta falta de criterio, de nuevo pudo haber pesado la diferencia de intereses y visiones que los participantes del panel tienen de la evaluación de riesgos profesionales (teórica vs. pragmática). Desde el punto de vista legal, práctico y estrictamente preventivo, está claro que no son aptos los cuestionarios cuya finalidad no sea identificar condiciones de trabajo peligrosas para la salud de los trabajadores (ITSS, 2012). Con esta realidad están muy familiarizados los técnicos de prevención de riesgos, pues la LPRL sólo contempla como medidas preventivas eliminar o reducir los factores de riesgo adversos. Sin embargo, existe una corriente académica que sobrepasa las competencias de los técnicos de prevención y por lo tanto fuera del marco de la LPRL, la *Psicología positiva*, que se basa en fomentar el bienestar y la salud de los trabajadores incluyendo evaluaciones de otros tópicos, como por ejemplo el *Engagement* (Carrasco et al., 2010, Pocinho y Perestrelo, 2011; Salanova y Schaufeli, 2009; Salanova et al., 2000; Ten Brummelhuis et al., 2012). Estos enfoques positivistas sin duda son beneficiosos para la salud, pero no deben mezclarse con las medidas legales para prevenir riesgos, en tal caso, y como bien dice la ITSS (2012) pueden utilizarse de complemento. Esta misma conclusión fue debatida con anterioridad por Niño (2006), quien retrata extensamente los errores que se cometen al evaluar riesgos psicosociales y que denomina “la ceremonia de la confusión” (p. 30). La existencia de confusión en este campo es evidente cuando conviven instrumentos de evaluación de riesgos psicosociales con medidores de clima laboral, de satisfacción o de estrés que sobrepasan los preceptos de la LPRL (véase el epígrafe 2.4 de esta tesis).

Sobre si es necesario tener conocimientos psicométricos para interpretar los resultados de las evaluaciones (ítem 25), no hubo acuerdo ni en contra ni a favor. Algunas voces sí reconocen la falta de formación específica de muchos profesionales que se dedican a la prevención, mencionando explícitamente entre las disciplinas necesarias para evaluar, la psicometría y la psicología del trabajo, entre otras (Meliá, 2008). Una vez más podría haber tenido influencia la heterogeneidad de la muestra, y sería de interés la detección de subgrupos, pues psicólogos y técnicos de prevención no cuentan con la misma formación en psicometría.

4.5.3. Implicaciones teóricas y prácticas

El eje central de esta investigación trataba de conocer las distintas posturas de los expertos en relación con los puntos controvertidos que rodean las evaluaciones de riesgo psicosocial. En base a ese planteamiento, la aportación más relevante ha sido profundizar en el conocimiento de los factores que dificultan dichas evaluaciones. El mero hecho de poder identificar los obstáculos es un paso fundamental para poder atajarlos, y en ese sentido la presente investigación ha ampliado el campo de estudio, al destapar varios puntos clave sobre los que sería necesario investigar y actuar.

En consecuencia, las principales implicaciones teóricas derivadas de este estudio enlazan con las limitaciones señaladas por los expertos como amenazas a las evaluaciones de riesgo psicosocial, destacando: (1) la confirmación del papel secundario que juegan estas evaluaciones a nivel organizacional, donde la escasa cultura preventiva tiene mucho que ver en la escasa sensibilización por los riesgos psicosociales, cuestión que coincide con los postulados de Cifre et al. (2011), Leka y Jain (2010), (ESENER-2 (2015) y con el informe de la ITSS (2015), entre otros trabajos; (2) la fusión de distintos intereses, como el académico, legal o intervencionista, propicia la confusión en cuanto al contenido de la evaluación, pues como indica la ITSS (2012) a veces se amplía a terrenos que se extralimitan del marco de la LPRL; (3) la necesidad de contar con la participación de los trabajadores, que aparte de ser una reiterada recomendación (ITSS, 2012; NTP 450 de Oncins y Almodóvar, 1997), no hay que olvidar que es una obligación recogida en la LPRL (arts.18.2, 34 y 36.2); (4) la reiterada ausencia de control de sesgos, a pesar de su transcendental influencia en la validez de los resultados (Anguiano-Carrasco et al., 2013) sigue sin prestársele una debida atención (Dupuis et al., 2015), asunto que contribuye a ensombrecer la credibilidad de las evaluaciones de riesgo psicosocial; y (5) la falta de seguimiento de un protocolo completo que oriente sobre las condiciones de administración de los cuestionarios en aras a minimizar posibles sesgos, como se recomienda desde AERA, et al. (2014) e ITC (2010).

En relación con las implicaciones prácticas, los resultados de esta investigación, al ser fruto del conocimiento y experiencia de los panelistas, han resultado ser una valiosa fuente de información que pueden servir de orientación para la mejora de los procesos de evaluación de riesgos psicosociales, ya que el simple conocimiento de las causas que dificultan los mecanismos de evaluación supone una oportunidad de avanzar en soluciones para amortiguarlas. En consecuencia, en la tabla 51 se presentan a modo de directrices prácticas, y de forma resumida, las principales recomendaciones extraídas del juicio experto que podrán contribuir a superar las amenazas identificadas, mejorando así la utilidad práctica del presente trabajo.

Tabla 51: Principales procedimientos y directrices sobre la evaluación de riesgos psicosociales
Fuente: Elaboración propia

Criterios que debería contener la herramienta de recogida de información para ser lo más objetivos posible:
<ul style="list-style-type: none"> - Los instrumentos deberán estar adaptados a la población a la que se dirigen (ej., tener en cuenta en empresas donde se encuentran trabajadores inmigrantes, discapacitados, etc.). - Se debe tener en cuenta el entorno y al trabajador como miembro de una organización, no como entes aislados.
Condiciones de administración:
<ul style="list-style-type: none"> - Debe ajustarse al contexto de evaluación. - Es necesario fomentar la participación y la autoconfrontación.
Diseño de instrumentos de medida:
<ul style="list-style-type: none"> - Incrementar la validez aparente de los instrumentos colaborando, para ello, con los profesionales de la PRL.
Aplicación:
<ul style="list-style-type: none"> - Administrar distintas pruebas que posibiliten la triangulación metodológica. - Antes de administrar un cuestionario es necesario explicar convenientemente los objetivos del mismo a los trabajadores. - Se debe garantizar la confidencialidad y el anonimato de la información obtenida. - Es necesario implicar a la dirección de la empresa en el proceso de la evaluación y que ésta indique que tras la evaluación se intervendrá según los recursos y prioridades que se establezcan. - Debe tomarse en consideración el contexto social de la empresa. - Se debe garantizar la credibilidad de la(s) persona(s) que lleve(n) a cabo la evaluación de riesgos psicosociales.
Sugerencia final:
<ul style="list-style-type: none"> - Es imprescindible complementar la información recabada con cuestionarios con otros métodos de obtención de información que permitan profundizar en las condiciones y entorno de trabajo.

Los expertos consultados han evidenciado la necesidad de contar con un protocolo completo y específico para evaluar riesgos psicosociales con medidas de autoinforme. Para llevarlo a cabo sería de gran utilidad desarrollar todas y cada una de las recomendaciones recogidas en la tabla 51, que se recuerda son fruto del consenso entre los mismos expertos. Asimismo, no deberían faltar en dicho protocolo las directrices indicadas en AERA et al. (2014) e ITC (2010), que recogen aspectos tan importantes como la preparación de los responsables de administrar los test, quienes no sólo deberían tener conocimientos metodológicos en cuestiones de identificación y reducción de posibles fuentes de error, sino que también deberían aprender o entrenar habilidades sociales como empatía o comunicación, entre muchas otras, destinadas por ejemplo a rebajar la ansiedad de los encuestados y a propiciar una mayor comprensión de los ítems incluidos en el cuestionario. Igual de importante es la inclusión de aspectos relacionados con el acondicionamiento o preparación del lugar destinado a rellenar los test de autoinforme, para así eliminar posibles distracciones e incomodidades que puedan influir en la calidad de las respuestas.

Los hallazgos derivados de este estudio deberían inducir a todos los implicados en el proceso de evaluación psicosocial a colaborar juntos (ej., instituciones oficiales, académicos, profesionales del sector o directivos). De ese modo, se salvaría el actual conflicto de intereses y podría profundizarse desde distintos enfoques y con mayor efectividad en cada una de las

limitaciones apuntadas por los expertos. Como consecuencia, podría conseguirse mejorar la calidad y cantidad de las evaluaciones, lo que a su vez podría contribuir en buena medida a rebajar las severas estadísticas relacionadas con el impacto del estrés laboral.

Por su parte, las organizaciones y/o instituciones oficiales deberían enfocar sus actuaciones en intentar solventar los obstáculos apuntados por los expertos. Por ejemplo, (1) se podrían orientar programas de investigación que indaguen en la influencia de los sesgos de respuesta en las medidas de autoinforme destinadas a evaluar riesgos psicosociales, pues no existen evidencias concluyentes sobre sus efectos; (2) podrían confeccionarse acciones formativas destinados a suplir las carencias en la identificación y control de sesgos, pues en consonancia con los resultados de este estudio ha quedado patente que los expertos académicos y profesionales no le dan la misma importancia, siendo precisamente el tema de los sesgos el que mayor controversia ha generado; (3) sería importante contar con programas de mayor sensibilización ante los riesgos psicosociales, pues el 100% de participantes en el estudio confirma que existe una acusada dejadez por parte de los directivos de las empresas a la hora de gestionar estos riesgos; y (4) también sería útil la creación de un espacio para la sinergia, donde todos los intereses relacionados con la prevención de riesgos psicosociales tuviesen cabida, aunando esfuerzos y experiencias, libre de recelos y competencias.

En resumen, los resultados de esta investigación han arrojado luz sobre las circunstancias que más lastran las evaluaciones de riesgo psicosocial. Actuar sobre ellas podrá poner fin al papel secundario que juega la esfera psicosocial en el ámbito preventivo.

4.5.4.Limitaciones y líneas de investigación futuras

A pesar del rigor con el que se ha diseñado y llevado a cabo el presente estudio, éste no carece de limitaciones. Lo que primero podría ser objeto de crítica es la consabida subjetividad en la que está basada la metodología (i.e., el juicio subjetivo de teóricos expertos), lo que llevaría a cuestionar el grado de conocimiento real de los expertos en el panel Delphi. Esta debilidad dependerá del acierto y destreza en la elección de panelistas. Para salvar esta limitación, se ha tratado con especial cuidado el proceso de selección de participantes aplicando criterios de inclusión rigurosos, como el *coeficiente K* (véase metodología, epígrafe 4.3.2). Para garantizar la calidad del estudio no se ha compartido ni información ni opiniones preconcebidas. Además de ello, se han estudiado en profundidad y con el mismo rigor metodológico tanto las convergencias como las divergencias.

En lo que respecta al tamaño de la muestra (16 panelistas), aunque se encuentra dentro del rango de 7 a 50 participantes, que es el tamaño considerado óptimo por los expertos para aplicar el método Delphi (Astarriaga, 2000; Delbecq et al., 1975; Landeta, 2002; J. Ruiz y Ispizua, 1989), lo cierto es que una muestra mayor hubiera permitido, estadísticamente, ir más allá de establecer el grado de consenso. En otras palabras, aunque el tamaño muestral conseguido no limita de ningún modo los resultados obtenidos, no permite aprovechar toda la información proporcionada por los expertos y explorar otras cuestiones ajenas al consenso en sí. Por ejemplo, una muestra mayor hubiera permitido comprobar la existencia de subgrupos, así como examinar si las divergencias aparecidas han sido fruto de la diferencia de intereses en la población seleccionada. Por otra parte, un número mayor de participantes podría contribuir a engrandecer con más puntos de vista las recomendaciones prácticas. Sería interesante que

futuros estudios pudiesen replicar esta investigación en muestras más amplias que permitan realizar análisis segmentados y comparaciones entre grupos con la debida potencia estadística. De ese modo se podría discernir si las diferencias de criterio de los expertos son debidas por ejemplo a sus distintos intereses o disciplinas formativas.

Otra posible salvedad es la delimitación geográfica del estudio al territorio nacional, lo que obliga a ser cautos a la hora de generalizar los resultados a otros países. Resultaría acertado ampliar esta investigación a expertos pertenecientes a otros estados de la UE, ya que así se lograría comparar los resultados geográficamente y daría la posibilidad de compartir experiencias y extraer recomendaciones de países con cultura preventiva más avanzada.

A pesar de sus limitaciones, los resultados de este estudio son sumamente valiosos, al tener como fuente a un grupo de profesionales y académicos, expertos conocedores del escenario real que encierra la prevención y evaluación de riesgos psicosociales de origen laboral.



SEGUNDO ESTUDIO





5. INVESTIGACIÓN CUANTITATIVA

5.1. INTRODUCCIÓN

5.1.1. Deseabilidad social como distorsión de las respuestas

De todas las proposiciones planteadas en el anterior estudio cualitativo, las relativas a los sesgos de respuesta han sido las que más controversia han suscitado. La falta de un posicionamiento claro por parte de los expertos ha impedido arrojar luz sobre si este tipo de sesgos influye en las puntuaciones de los test que evalúan riesgos psicosociales. En consecuencia, las posturas sobre la DS tampoco han quedado claras, a pesar de ser el sesgo más estudiado por la comunidad científica (Cosentino y Castro, 2008), y de existir numerosas evidencias de su papel distorsionador en los resultados de evaluaciones de personalidad medidas con autoinforme; repercutiendo sobre todo en la práctica clínica, forense y médico-legal (González e Iruarrizaga, 2005). También se han demostrado similares efectos en el ámbito de la psicología del trabajo y de las organizaciones, por ejemplo, en los cuestionarios empleados en selección de personal (Salgado, 2005). Como se ha sugerido anteriormente, Semmer et al. (2004) reconocen que los sesgos de respuesta y la afectividad negativa pueden inflar las correlaciones entre bienestar laboral y estresores, incluso llegar a generar asociaciones espurias. De hecho, a raíz de los altos índices de estrés laboral extraídos de algunas encuestas en los periodos 1997-2006 (ECVT; ENS; ISPP), una de las interpretaciones realizadas ha puesto en escena la DS como factor responsable de inflar esas puntuaciones (Gamero-Burón y González, 2013), al indicar que los trabajadores han podido responder guiados por los estereotipos sociales. De ese modo, mostrar que se sufre estrés implicaría ser un miembro valioso de la empresa. Duro (2005) se suma a estos argumentos afirmando que los autoinformes que miden variables relacionadas con el ámbito psicosocial laboral son sensibles a la DS, sin embargo, va más allá, y asegura que dependiendo de lo que evalúen los ítems, la distorsión será positiva o negativa. Esto es, si se evalúan variables que recogen consecuencias laborales negativas para el trabajador como el absentismo, la tendencia de respuesta de los trabajadores es marcar las puntuaciones que reflejen conductas menos absentistas, sin embargo, cuando lo que se mide implica actuaciones positivas del trabajador, como la capacidad personal, la tendencia de respuesta es marcar las opciones que contienen las puntuaciones que evidencian mayor valor del trabajador como activo de la empresa.

Sin embargo, en lo referente a cuestionarios destinados a evaluar condiciones de trabajo no se han encontrado estudios concluyentes sobre los efectos de la DS, convirtiéndose en una incógnita que convendría despejar por su gran utilidad, tanto teórica como práctica, para todos los implicados en el sector de la prevención. Explorar las herramientas de evaluación de riesgos psicosociales podría desvelar si son susceptibles a sesgos de respuesta, y si es así, podrían hacerse esfuerzos por conocer si algunos factores son más sensibles a la DS que otros (ej., demandas cuantitativas, conflicto de rol, o doble presencia).

Como se ha expuesto en el marco conceptual referente a los sesgos de respuesta, coexisten una variedad de técnicas para controlar la DS, desde las más sencillas como escalas, hasta las más complejas como la técnica basada en análisis factorial de Ferrando (2005) y Ferrando y Anguiano-Carrasco (2009, 2010). Esta última es la que se ha seleccionado para identificar y depurar la DS, ya que permite extraer medidas puras de DS sin tener que eliminar información.

5.1.1. Afecto negativo como variable de confusión

Resulta evidente que ante la existencia de un factor de riesgo, no todos los trabajadores enferman. Sino que dependerá del nivel de exposición, del tiempo expuesto y de las características personales. Además, cuando el factor de riesgo es un estresor, muchos autores confirman que la personalidad es una variable que interviene determinadamente en su percepción (Caicoya, 2004; J. Martín et al., 2007; Staw, Bell y Clausen, 1986). Desde ese punto de vista son muchas las investigaciones que han mostrado correlaciones significativas entre la personalidad y el bienestar laboral (Brunborg, 2008; Judge, Heller y Mount, 2002). Dentro de los rasgos más investigados destacan el afecto negativo y positivo (Man, 2013; Mejía, Laca y Gondra, 2009; Warr, 1996, 1998; Weiss y Cropanzano, 1996). Por consiguiente, aparte de los sesgos de respuesta, la afectividad podría entrar en juego como posible fuente de distorsión (Caicoya, 2004, F. Jones y Kinman, 2001; Razavi, 2001).

Una definición clara de *afecto negativo* (AN) es recogida por Watson y Clark (1984, citado por la EU-OSHA, 2005):

Un rasgo de la personalidad general que refleja las diferencias individuales en la emotividad negativa y en el concepto de uno mismo, es decir, concentrarse sobre los aspectos negativos de cada cosa y experimentar una dosis considerable de distrés en todas las situaciones (p. 57).

El AN refleja un estado emocional que en niveles altos se caracteriza por variedad de estados de ánimo aversivos, que incluyen miedo, disgusto, ira, culpa y nerviosismo, mientras que el bajo AN es un estado de calma y serenidad (Watson, Clark y Tellegen, 1988). El *afecto positivo* (AP) se refiere a una dimensión en la que niveles altos se caracterizan por alta energía, concentración completa, agradable dedicación y sentimientos de logro, mientras que el bajo AP se caracteriza por tristeza y letargo. Estas dos dimensiones son catalogadas como bipolares, aunque en la práctica solo cobra importancia el polo alto de ambas, ya que los polos bajos representan la ausencia relativa de implicación afectiva (Sandín et al., 1999; Watson y Tellegen, 1985). Tradicionalmente se asume que la extroversión está fuertemente asociada con el AP, siendo el neuroticismo la principal fuente de AN (Beer, Watson y McDade-Montez, 2013; Connolly y Viswesvaran, 2000; Costa y McCrae, 1980; Judge et al., 2002; Rusting y Larsen, 1998).

Los sujetos que presentan AN alto poseen una visión más negativa, tanto de ellos mismos como del mundo que los rodea, percibiendo mayor amenaza y daño potencial (J. Martín et al., 2007). Además, presentan mayores actitudes negativas, quejas somáticas problemas de salud y trastornos psicofisiológicos (Clark y Watson, 1988; Robles y Páez, 2003). Especialmente relacionado con el contexto laboral, los trabajadores dominados por este rasgo consideran los

factores psicosociales de manera más desfavorable (J. Martín et al., 2007), y cuentan con mayor probabilidad de sentirse tensos al percibir experiencias de estrés negativo más recurrentes y de mayor intensidad (Burke, 2001; Magnus, Diener, Fujita y Pavot, 1993; Sandi y Calés, 2000). De hecho, se han hallado correlatos significativos entre alta AN y percepciones estresantes de las condiciones de trabajo (Brief et al., 1988; Elliot, Chartrand y Harkins, 1994; Levin y Stokes, 1989). Otro de los aspectos con los que se relaciona el AN es con la percepción de la salud (Heinisch y Jex; 1997; González-Gutiérrez et al., 2009; Warr, 2007; Watson y Pennebaker, 1989; Zautra et al., 2005). Así, Harding (1982) halló relación significativa entre el AN y las medidas de autoinforme sobre salud, ansiedad y preocupaciones, encontrándose como mejores predictores el estatus socioeconómico y la competencia personal.

El AN explica la varianza hallada en autoinformes de medida de muchas variables, entre ellas el estrés, estresores o síntomas psicosomáticos, inflando ampliamente las correlaciones entre el estresor y la variable dependiente, hecho que pone en tela de juicio la validez de las medidas aplicadas mediante autoinforme (Brief et al., 1988; Watson y Pennebaker, 1989), pues podría implicar “un problema de confusión o circularidad implícita de medidas entre predictor y criterio a causa de una «tercera variable» subyacente” (Santed, Sandín, Chorot, Olmedo y Campayo, 2001, p. 42). Esto hace suponer que los trabajadores con un perfil de personalidad dominado por el rasgo de AN puntuarán más alto que los trabajadores con bajo AN en medidas de autoinforme que evalúen condiciones laborales adversas. Al respecto, EU-OSHA (2005) es una más en señalar que el AN podría actuar como variable de confusión pues no sólo estaría influyendo en la percepción que los trabajadores tienen sobre el ambiente laboral, sino que también podría intervenir en la valoración de su propio estado de salud o de bienestar, hecho que explicaría gran cantidad de correlaciones entre los riesgos percibidos y los resultados manifestados.

En ese orden de ideas, EU-OSHA (2005) menciona la existencia de distintas posturas acerca del papel distorsionador del AN o el método común de la varianza en la relación estrés-tensión. Y como estrategia de control parcial proponen la triangulación, que consiste en conseguir datos cualitativos de distintas fuentes o conjugar datos cuantitativos con datos cualitativos, tal y como recomiendan, entre otros, autores como Caicoya (2004), Cox y Ferguson (1994), Cox et al. (2000) y F. Jones y Kinman (2001). Concretamente proponen recoger datos de tres esferas distintas (Cox et al., 2000): (i) datos objetivos y subjetivos sobre ambiente laboral, tanto físico como psicológico; (ii) autoevaluación de la propia experiencia del estrés como emoción negativa; e (iii) identificación de posibles alteraciones de su comportamiento, fisiología o estado de salud que puedan relacionarse con los datos sobre experiencias de estrés recabados en los dos puntos anteriores. Estos últimos datos pueden recabarse mediante medidas denominadas no intrusivas, Del Hoyo (2004) propone las siguientes como signos de existencia de estrés: (1) Disminución de la producción (calidad, cantidad o ambas); (2) falta de cooperación entre compañeros; (3) aumento de peticiones de cambio de puesto de trabajo; (4) necesidad de una mayor supervisión del personal; (5) aumento de quejas en los clientes; (6) empeoramiento de las relaciones humanas; (7) falta de orden y limpieza; (8) aumento del absentismo; (9) aumento de incidentes y accidentes; (10) aumento de quejas al servicio médico; y (11) aumento en el consumo de tabaco, alcohol y medicamentos.

Para seguir esta estrategia de la triangulación se requiere utilizar simultáneamente varios instrumentos, como entrevistas individuales, cuestionarios, observación, grupos focales, consultas de datos estadísticos de presencia de factores de riesgo psicosocial (indicadores de absentismo, incapacidad temporal, accidentes laborales, enfermedades profesionales, quejas, rendimiento, etc.), lo cual implica costes económicos y temporales elevados. Precisamente, resulta paradójico que éstos sean los dos principales escollos para abordar la prevención de riesgos en las empresas europeas según la encuesta ESENER-2, donde un 25.50% de centros de trabajo europeos achacan como principal dificultad la falta de tiempo o personal y un 22.60% la falta de presupuesto (Almodóvar et al., 2015).

Otra de las soluciones aportadas y ampliamente aceptada, aparte de la triangulación, es la recomendación de incluir medidas de AN en los autoinformes para poder controlar estadísticamente el papel que juegan en las investigaciones, sobre todo si se van a medir variables de salud subjetiva y/o estresores (ej., Parkes, Mendham y von Rabenau, 1994; Moyle, 1999; Wofford, Goodwin y Daly, 1999, citados en Razavi, 2010). Esta cuestión es objeto de un debate recientemente reavivado (Razavi, 2010), pues los correlatos no se han encontrado de forma sistemática, sino que hay investigaciones en las que tal efecto no ha sido observado (García-Layunta, Tomás y Oliver, 2005).

5.2. JUSTIFICACIÓN DEL PROBLEMA, OBJETIVOS E HIPÓTESIS DE LA INVESTIGACIÓN

De acuerdo con la revisión teórica realizada, con los resultados y conclusiones obtenidos del estudio cualitativo, y ante la falta de estudios empíricos que muestren evidencias sobre el papel distorsionador de la DS y AN, se justifica un segundo estudio cuantitativo que profundice sobre la influencia de estas variables en las medidas de evaluación de riesgos psicosociales. De este modo se pretende avanzar empíricamente en el ámbito de la prevención de riesgos psicosociales, más allá de los estudios meramente descriptivos de perfiles de riesgo. Es preciso emprender nuevas investigaciones exploratorias que abran camino y faciliten el desarrollo de futuros modelos e hipótesis más específicas.

5.2.1. Objetivo general

El presente trabajo pretende profundizar en el conocimiento de la posible influencia de sesgos de respuesta en los resultados de las evaluaciones de riesgos psicosociales de origen laboral obtenidas con medidas de autoinforme.

5.2.2. Objetivos específicos

- a) Identificar si la DS influye distorsionando las puntuaciones de los autoinformes que miden riesgos psicosociales en el trabajo.
- b) Analizar el funcionamiento diferencial del ítem para identificar si el rasgo de AN actúa desvirtuando los resultados e inferencias realizadas de los cuestionarios de evaluación de riesgos psicosociales.

5.2.3. Hipótesis

El panorama descrito a la luz del juicio experto, sobre la influencia de los sesgos de respuesta en las autoevaluaciones de riesgos psicosociales (reflejado en el estudio Delphi),

sumado a las evidencias de estos sesgos en otras medidas autoinformadas, hace prever los siguientes resultados presentados en forma de hipótesis de trabajo:

H1: La medida de los factores psicosociales de riesgo estará sesgada debido al efecto de la discapacidad social.

H2: La medida de los factores psicosociales de riesgo contiene ítems con DIF en contra de los trabajadores que presentan alto afecto negativo.

5.3. METODOLOGÍA DE ESTUDIO

5.3.1. Selección de herramienta de evaluación de riesgos psicosociales

Para seleccionar la herramienta de evaluación de riesgos psicosociales más idónea a utilizar en esta investigación, se han seguido los criterios establecidos por Moreno y Báez (2010, para un resumen véase tabla 21 en el epígrafe 2.4). Según los argumentos de estos autores, solamente dos de las herramientas revisadas cumplirían con la mayor parte de los criterios establecidos, el CoPsoQ-ISTAS 21 y el F-PSICO. Tras una revisión más detenida sobre sendos instrumentos, se ha encontrado que el análisis de las propiedades psicométricas del F-PSICO, herramienta elaborada por el INSHT (Martín-Daza y Pérez, 1997), fue realizado en el año 2010 (Ferrer et al., 2011), y su estudio versa sobre los 86 ítems de las subescalas. Sin embargo, la nueva versión (versión 3.1) es del año 2014 y contiene varias modificaciones respecto a la versión anterior, como el aumento de ítems (89), por lo que el estudio de las propiedades psicométricas publicado no puede reflejar fielmente las propiedades de la nueva versión. Por otra parte, Peiró et al. (2013) afirman que la evidencia empírica publicada es limitada y no correlaciona suficientemente con las variables sobre salud, en la misma línea existen varios estudios donde la consistencia interna refleja valores pobres (ej., Callejón-Ferre et al., 2015; Meseguer et al., 2007).

En lo que respecta al CoPsoQ, ha sido desarrollado inicialmente en Dinamarca en 1997 (CoPsoQ I). Fue validado y estandarizado en una amplia muestra de trabajadores daneses (Kristensen, Hannerz, Høgh, y Borg, 2005), para posteriormente validarse y aplicarse en numerosos países. Prueba de su gran difusión es que se encuentra disponible en más de 25 idiomas (Berthelsen, Hakanen, Kristensen, Lönnblad y Westerlund, 2016; OIT, 2016). En España ha sido adaptado por el Instituto Sindical de Trabajo, Ambiente y Salud (ISTAS) con el nombre de CoPsoQ-ISTAS 21, y ha sido estandarizado y validado en una amplia muestra representativa de población española asalariada, siendo la única herramienta que ofrece valores de referencia para cada una de las dimensiones (Moncada et al., 2008). Valores que a su vez son actualizados, al menos, una vez cada cinco años (Moncada, Utzet et al. 2014). Es una de las herramientas más utilizadas y, a su vez, tiene el respaldo del sindicato UGT (Moreno y Báez, 2010). Además de ello, cuenta con numerosa literatura científica (Kristensen et al., 2005), como ejemplo los más de 100 artículos publicados en revistas internacionales (Moncada, Utzet et al., 2014). En el 2010 se publicó la segunda versión en Dinamarca (Pejtersen, et al., 2010) y en el 2014 la versión española (Moncada, Llorens et al., 2014). La utilización de este cuestionario en tantos países ha contribuido a su validación. De ese modo pueden encontrarse numerosas evaluaciones de riesgos psicosociales que al utilizar el CoPsoQ han ratificado sus propiedades psicométricas. La mayor parte valida su consistencia interna, su fiabilidad y/o su validez de

criterio, siendo mucha más escasa la información referida a la validez de contenido (Berthelsen et al., 2016).

En función a estos argumentos se ha descartado la utilización del F-PSICO, optando por emplear la versión corta del cuestionario CoPsoQ-Istas21 versión 2 (Moncada, Llorens et al., 2014) compuesta por 30 ítems. Si se compara con los 89 ítems del F-PSCICO, se ofrece un motivo más de preferencia por el CoPsoQ, debido a que contribuye en mayor medida a minimizar los sesgos de fatiga (Choi et al., 2010; Robins, Hendin y Trzesniewski., 2001), y de repuesta invariable atribuibles a medidas largas. Precisamente, los más de 100 ítems que constituyen la versión larga del CoPsoQ-Istas 21, específico para investigación, ha sido el principal motivo de su descarte.

La adaptación del CoPsoQ II a la realidad cultural, lingüística y laboral española ha sido publicada por Moncada, Utzet et al. (2014) a partir de los siguientes estudios: (i) trabajo de campo de Moncada, Utzet, Llorens, Galtés y Moreno (2010) basado en una muestra representativa de la población asalariada española y estratificada por áreas geográficas (N= 5110), y (ii) descripción sobre las desigualdades psicosociales de origen laboral entre los asalariados de España y Dinamarca correlacionadas con el estatus socioeconómico (Moncada, Pejtersen et al., 2010).

Según los datos publicados por Moncada, Utzet et al. (2014), la validez factorial de las escalas se verificó en la versión larga del cuestionario mediante análisis exploratorio utilizando el Criterio Oblimin (método de rotación oblicua) y el método de extracción factorial de máxima verosimilitud. La versión media y corta son desarrolladas a partir de la versión larga. En términos generales la consistencia interna de las escalas de la versión media es alta. De las 24 dimensiones que posee el cuestionario, 19 presentan valores alfa de Cronbach por encima .79, el valor más bajo es de .59 y corresponde a la dimensión *de previsibilidad*.

La selección de ítems se realizó mediante correlaciones inter-ítem, e ítem total, observando el efecto de alfa de Cronbach eliminado cada ítem; los valores de las correlaciones considerados óptimos se situaron ente los valores .20 y .50, pues correlaciones mayores indicarían redundancia de la medida y los menores supondría que los ítems no forman parte de la misma dimensión conceptual. Asimismo, los índices kappa entre la versión media y la versión corta resultó satisfactoria para todas las dimensiones con valores que van desde .62 a .84.

Se realizó una búsqueda específica acerca de la estructura factorial del cuestionario, a saber: matrices factoriales y/o valores de saturación de los ítems en cada factor. Sin embargo, no se encontró información adicional, ni sobre la versión original danesa, ni sobre la versión adaptada a la población asalariada española, a pesar de consultar en las páginas web oficiales de CoPsoQ e ISTAS 21, y de indagar en las bases de datos PsycINFO , SCOPUS, REBIUN, ProQuest, SciFinder y en Google Scholar. Del mismo resultado infructuoso hacen eco Moreno y Báez (2010) en su estudio sobre evaluación de riesgos psicosociales.

5.3.2.Diseño y Procedimiento de recogida de datos

Conforme a los objetivos de esta investigación, que están centrados en profundizar sobre la influencia de los sesgos de respuesta en los resultados de las evaluaciones de riesgos psicosociales, se plantea este estudio desde un estudio de campo cuantitativo de corte

exploratorio. Para la realización del presente estudio se ha confeccionado un cuestionario autoadministrado, utilizando la aplicación Formularios de Google, la cual permite, además, la posterior recogida de datos on-line. Una copia del mismo se muestra en el Anexo II. La mayor parte de los ítems del cuestionario (a excepción de las variables sociodemográficas y sociolaborales) fueron diseñados con la opción de respuesta obligatoria con objeto de minimizar los valores perdidos; es decir, si los cuestionarios contenían preguntas sin contestar automáticamente eran eliminados de la muestra.

El cuestionario ha sido estructurado en varias partes, a saber: (1) la correspondiente a la presentación del estudio y sus objetivos, donde a su vez se solicita la participación voluntaria y se resalta la importancia de responder con sinceridad, así como también se garantiza la confidencialidad y el anonimato de la información; (2) el apartado referente a las variables sociodemográficas de edad, sexo, nacionalidad, situación laboral, puesto de trabajo y otras variables de interés, constituyendo un total de 10 ítems; y (3) la que recoge los constructos objeto de investigación; en concreto, exposición a riesgos psicosociales, deseabilidad social y afecto, que comprenden 54 ítems. Por último, se añadió un ítem adicional para observaciones, comentarios y/o sugerencias. En total el cuestionario está conformado por 65 ítems.

La recogida de datos se realizó entre el 27 de junio y el 9 de septiembre del 2016, mediante la plataforma on-line anteriormente mencionada. Uno de los motivos de haber elegido este procedimiento fue que permite conseguir de forma rápida muestras heterogéneas muy amplias (e.g., diferentes edades, sectores empresariales, origen geográfico, puestos de trabajo, entre otras variables). Además de ello y según diversos estudios, las garantías metodológicas de la investigación a través de datos extraídos online son similares al procedimiento presencial o de investigación tradicional (Buchanan y Smith, 1999; Gosling, Vazire, Srivastava y John; 2004, Schleyer y Forrest, 2000). Sin embargo, cada vez son más los investigadores que recomiendan este método en aras a disminuir las respuestas socialmente deseables, pues la separación física entre la persona encargada de administrar el test y el informante contribuye al anonimato (Kreuter, Presser y Tourangeau, 2008; Moorman y Podsakoff, 1992; Pettit, 2002; Risko et al., 2006).

Dicha aplicación de Google genera un enlace automático donde se aloja el cuestionario, el *link* fue compartido a través del correo electrónico y de la red social para profesionales *LinkedIn*, con objeto de conseguir la mayor muestra posible. Una vez contestados todos los ítems por parte de los encuestados, éstos solo tenían que pulsar la opción “enviar” y automáticamente los datos eran recibidos en una plantilla electrónica en forma de cálculo, donde es imposible conocer la autoría de cada cuestionario, hecho que refuerza el carácter anónimo de la investigación. A su vez, todos los participantes tuvieron acceso mediante e-mail o mediante la plataforma social *LinkedIn* para ponerse en contacto con la autora y poder plantear las cuestiones o aclaraciones que considerasen pertinentes, como así hizo una gran parte de los encuestados cuando manifestaron su interés y apoyo en el estudio (aproximadamente, un 37% de los participantes)

Además, esta investigación se ha realizado siguiendo las directrices éticas marcadas por el Código de conducta para investigaciones del psicólogo de la APA (2010), a saber: informar a los participantes en el estudio del propósito de la investigación y de la duración aproximada, garantizar el anonimato y la confidencialidad de los datos, facilitar un espacio y dirección de

contacto para responder dudas o solicitar información adicional, así como informar al término de la investigación (de modo general) de los principales resultados y conclusiones.

5.3.3. Depuración de datos

La matriz de datos bruta estuvo formada por 642 casos y se sometió a un proceso de depuración antes de tratarla estadísticamente con objeto de detectar y corregir posibles errores. Las pruebas de exploración llevada a cabo siguieron las fases mostradas en la tabla 52:

Tabla 52: Fases de la depuración de datos
Fuente: Elaboración propia

ORDEN	A. ESTIMACIÓN DE ERRORES DE GRABACIÓN:
1º	Revisión de un 10% aleatorio de los casos introducidos en el SPSS
2º	Detección de errores comprobando las discrepancias entre los datos recogidos en la matriz bruta (Excel) y los datos importados al SPSS
3º	Estimación del porcentaje de error contenido en toda la matriz de datos: $(\text{Errores detectados} [\text{número total de cuestionarios} \times \text{número total de variables}]) \times 100$
ORDEN	B. ANÁLISIS EXHAUSTIVO DE:
1º	Valores perdidos
2º	Valores fuera de los rangos admitidos en la variable

Siguiendo estos criterios se revisaron 70 casos seleccionados aleatoriamente, entre los cuales se detectaron 11 datos erróneos, cifra que constituye un 0.02% de porcentaje de error. En cuanto a los valores perdidos, sólo fueron observados en las variables sociodemográficas y/o sociolaborales, pues tal y como se ha mencionado en párrafos precedentes, se diseñó el proceso de recogida de datos de tal forma, que los cuestionarios con respuestas incompletas en el apartado de las variables de interés no eran guardados en la matriz de datos; a su vez, sólo era permitida una sola respuesta por ítem, de esta manera no hubo cuestionarios con respuestas nulas ni incompletos en la muestra seleccionada, y por consiguiente, *cero* valores perdidos en dichas variables. El siguiente análisis de la depuración de datos consistió en explorar los valores fuera de rango; sin embargo, otra de las ventajas del diseño del cuestionario on-line radica en que las distintas elecciones de respuesta son fijadas de antemano, por lo que no cabe la posibilidad de responder una opción distinta a las indicadas, esto implica la ausencia de valores atípicos o fuera de los rangos admitidos en la variable.

5.3.4. Participantes

Para determinar la selección de la muestra se ha perseguido la mayor representatividad posible. Ante la dificultad de realizar un muestreo aleatorio, y debido a las oportunidades de acceso al universo, se optó por un muestreo incidental por cuotas balanceado por sexo, dada la semejanza con el muestro probabilístico estratificado. Para las cuotas se ha utilizado de referencia la distribución de ocupados agrupados por sexo de la Encuesta de Población Activa (EPA) publicada por el Instituto Nacional de Estadística (INE, 2016). Las cifras se exponen en la tabla 53 expresadas en miles.

Tabla 53: Muestreo en base al número de ocupados por grupo de sexo
Fuente: Elaboración propia

Población ocupados en España*			Muestra del estudio	
Sexo	N	Porcentaje	n	Porcentaje
Mujer	84 14,8 (miles)	45.42	254	45.1
Hombre	10 112,8 (miles)	54.58	309	54.9
Total	18 527,6 (miles)	100	563	100

* Datos de la encuesta EPA 2016 del tercer trimestre

La unidad de análisis está conformada por trabajadores en activo y por cuenta ajena² pertenecientes a ambos sexos y a cualquier sector de actividad en el territorio nacional. Se dio por finalizado el sondeo al alcanzar las cuotas determinadas teniendo en cuenta los siguientes criterios de exclusión: (1) Participantes en situación de desempleo (39 sujetos), que al no ser trabajadores por cuenta ajena no son objeto de evaluación de riesgos laborales; y (2) trabajadores por cuenta propia (40), por los mismos motivos que el anterior criterio, a excepción de los trabajadores autónomos económicamente dependientes (TRADE). En base a estos criterios se excluyeron de la muestra original 79 sujetos, quedando conformada la muestra final por 563 participantes.

Al ser el muestreo no probabilístico, no procede realizar el cálculo del margen de error ni el nivel de confianza del tamaño de la muestra; sin embargo, el margen de error máximo para una muestra aleatoria simple del mismo tamaño sería del 4.10% para un nivel de confianza del 99%, para un nivel de confianza del 95% el máximo error sería de 5.40%.

Según Nunnally (1978, citado por Morales, 2012a), cuando el estudio trate sobre el análisis de ítems, como es el caso que nos ocupa, el tamaño mínimo de la muestra recomendado es de 5 sujetos por ítem. Siguiendo dicho criterio, el tamaño de muestra mínimo para poder realizar análisis estables en esta investigación sería de 320 informantes, pues el cuestionario consta de 64 ítems incluyendo las variables demográficas y sin contar con el apartado de observaciones y comentarios. Para efectuar análisis factoriales las recomendaciones mínimas se establecen en muestras de al menos 200 informantes (Lloret-Segura, Ferreres-Traver, Hernández-Baeza y Tomás-Marco, 2014; Morales, 2012b). Por consiguiente, la muestra alcanzada en el trabajo de campo (n=563), al sobrepasar los requerimientos mínimos, cumple los estándares establecidos en la literatura psicométrica.

5.3.4.1. Distribución por sexo y edad

De los 563 participantes un 45.10% corresponde a mujeres y un 54.90% a hombres, sin existir valores perdidos. Estos datos hacen considerar la muestra equilibrada y homogénea, ya que las tasas son básicamente coincidentes con los datos de la última encuesta EPA (2016). En cuanto a la edad de los encuestados, se presenta en intervalos de 10 años, comprendida entre menores de 26 y mayores de 55 años, siendo la franja de mayor participación la que se encuentra

² Dentro de la población objeto de estudio no se han eliminado los trabajadores autónomos económicamente dependientes (TRADE) por estar dentro de este colectivo los denominados “falsos autónomos” o trabajadores por cuenta ajena encubiertos (Pérez-Amorós, 2015).

entre los 36 y 45 años (45.30%) y la menos frecuente es la relativa los menores de 26 años con un 3% de observaciones. No existen valores perdidos (ver tabla 54).

Tabla 54: Distribución en función del sexo y de la edad

Fuente: Elaboración propia

Sexo	n	Porcentaje	Edad	n	Porcentaje
Mujer	254	45.1	Menos de 26 años	17	3
Hombre	309	54.9	Entre 26 y 35 años	121	21.50
Total	563	100	Entre 36 y 45 años	255	45.30
			Entre 46 y 55 años	130	23.10
			Más de 55 años	40	7.10
			Total	563	100

5.3.4.2. Distribución en función de la nacionalidad

En la tabla 55 se presenta la frecuencia y porcentajes de la muestra, donde un 86.70% de participantes son de nacionalidad española y un 9.20% de nacionalidad distinta. Los valores perdidos recaen en 23 sujetos que no contestaron esta parte de la encuesta (4.10%). Según la EPA (INE, 2016), en el tercer trimestre, el 10.86% de trabajadores ocupados son de nacionalidad extranjera (2 012.2 miles).

Tabla 55: Distribución en función de la nacionalidad

Fuente: Elaboración propia

Países	n	Porcentaje
España	488	86.70
Argentina	8	1.40
Bélgica	1	0.20
Bolivia	2	0.40
Brasil	4	0.70
Chile	4	0.70
Colombia	3	0.50
Ecuador	1	0.20
Francia	1	0.20
Guatemala	2	0.40
Lituania	1	0.20
México	3	0.50
Perú	8	1.40
Portugal	6	1.10
Venezuela	8	1.40
Total	540	95.90
Perdidos	23	4.10
Total	563	100.00

5.3.4.3. Distribución según la situación laboral

Respecto a la situación laboral una amplia mayoría tiene contrato indefinido (67.10%), seguidos de los trabajadores con contrato temporal (27%). El menor porcentaje de la muestra

hace referencia a los becarios (1.10%). En este apartado del cuestionario contestaron todos los encuestados (ver tabla 56).

Tabla 56: Distribución según la situación laboral

Fuente: Elaboración propia

Situación laboral	n	Porcentaje
Contrato indefinido	378	67.10
Contrato temporal	152	27.00
Interino	5	0.90
TRADE	22	3.90
Becario	6	1.10
Total	563	100.00

5.3.4.4. Distribución según el puesto de trabajo que ocupan

Se han categorizado los puestos de trabajo siguiendo la Clasificación Nacional de Actividades Económicas (CNAE-2009) (ver tabla 57). Para adaptar dicha clasificación a las respuestas de los encuestados se han agregado 3 categorías adicionales correspondientes a (i) mandos intermedios, (ii) mandos superiores, y (iii) otros servicios. Solamente dos trabajadores han dejado en blanco esta parte del cuestionario y uno ha contestado indebidamente. Los participantes pertenecen a muy diferentes sectores, lo que hace que la muestra sea también suficientemente heterogénea.

Tabla 57: Distribución según el puesto de trabajo.

Fuente: Elaboración propia

Puesto de trabajo	n	Porcentaje	Puesto de trabajo	n	Porcentaje
Industria extractiva	1	0.20	Actividades sanitarias y servicios sociales	36	6.40
Industria manufacturera	2	0.40	Prevención	158	28.10
Actividades administrativas y servicios auxiliares	53	9.40	Actividades artísticas, recreativas y de entretenimiento	3	0.50
Construcción	7	1.20	Trabajadores del hogar	1	0.20
Comercio	22	3.90	Mandos intermedios	86	15.30
Hostelería	2	0.40	Mandos superiores	30	5.30
Información y comunicaciones	9	1.60	Otros servicios	4	0.70
Actividades financieras	5	0.90	No procede	1	0.20
Actividades profesionales, científicas y técnicas	84	14.90	Total	561	99.60
Administración pública	12	2.10	Perdidos	2	0.40
Educación	45	8.00	Total	563	100.00

5.3.4.5. Distribución según antigüedad en el puesto de trabajo y en la empresa

Tanto la antigüedad en el puesto de trabajo como en la empresa se ha dividido en 10 intervalos crecientes que van desde los 6 meses hasta los 20 años. El porcentaje de frecuencia más alto en ambas distribuciones se encuentra entre los 6 meses y un año, concretamente 13.70% de antigüedad en el puesto y 15.50% en la empresa. Es seguido muy de cerca por el

periodo de antigüedad comprendido entre un año y dos años, 13.10% y 12.80% respectivamente. Por otro lado, los porcentajes más bajos de frecuencia se refieren a trabajadores con más de 20 años en el mismo puesto (4.80%) y entre 6 y 8 años de permanencia en la misma empresa (6.20%). Los valores perdidos fueron de 8 casos para la antigüedad en el puesto y sólo uno para la antigüedad en la empresa (ver tabla 58).

Tabla 58: Distribución según la antigüedad en el puesto y en la empresa
Fuente: Elaboración propia

Intervalos	Antigüedad en el puesto		Antigüedad en la empresa	
	n	Porcentaje	n	Porcentaje
Menos de 6 meses	67	11.90	64	11.40
>6 meses a 1 año	77	13.70	87	15.50
> 1 año a 2 años	74	13.10	72	12.80
> 2 años a 4 años	57	10.10	53	9.40
> 4 años a 6 años	49	8.70	46	8.20
> 6 años a 8 años	37	6.60	35	6.20
> 8 años a 10 años	62	11.00	63	11.20
> 10 años a 15 años	72	12.80	58	10.30
> 15 años a 20 años	33	5.90	37	6.60
Más de 20 años	27	4.80	47	8.30
Total	555	98.60	562	99.80
Perdidos	8	1.40	1	0.20
Total	563	100.00	563	100.00

5.3.4.6. Distribución según tiempo de baja médica en los últimos 12 meses

Todos los participantes respondieron a este ítem, donde una gran mayoría (83.50%) indica no haber estado de baja en los últimos 12 meses, frente al 14.90% que sí lo ha estado. Dentro de estos últimos tan sólo un 1.60 % han declarado que la baja ha estado relacionada con el trabajo, tal y como se presenta en la tabla 59.

Tabla 59: Distribución en función de la baja médica
Fuente: Elaboración propia

Baja médica	n	Porcentaje
NO	470	83.50
SÍ	84	14.90
Baja relacionada con el trabajo	9	1.60
Total	563	100.00

5.3.5. Instrumentos

Como se apuntó anteriormente, el cuestionario elaborado para su cumplimentación por parte de los trabajadores que aceptaron participar en el estudio, consta de varias partes (Anexo II):

- Cuestiones sociodemográficas y sociolaborales (10 ítems).

- b) Escala CoPsoQ- ISTAS 21 v.2 (30 ítems).
- c) Ítems utilizados a modo de marcadores de deseabilidad social (4 ítems de control).
- d) Escala PANAS (Afecto negativo y Afecto positivo) (20 ítems).

5.3.5.1. Cuestionario de Evaluación de Riesgos Psicosociales en el Trabajo del método ISTAS21 (CoPsoQ) (versión 2)

El formato de respuesta de esta escala es de tipo Likert de cinco puntos (0 a 4). Midiendo 6 dimensiones principales, divididas a su vez en varias subdimensiones (Moncada, Llorens et al., 2014), tal y como se muestran a continuación:

- I. Exigencias psicológicas en el trabajo (Subdimensiones: Exigencias cuantitativas, ritmo de trabajo, exigencias emocionales) (6 ítems, $\alpha=.77$). Estas exigencias pueden ser cuantitativas, referidas a la relación entre el volumen de tareas y el tiempo para realizarlas (presión de tiempo), escapando al control del trabajador. No tener tiempo suficiente para llevar a cabo la gestión de las altas exigencias puede provocar tener que trabajar a un ritmo rápido; no ser capaz de llevar el trabajo al día o tener que acumular las tareas (Moncada, Llorens et al., 2014; Ohly y Fritz, 2010). También está relacionada con la distribución irregular de las tareas y con la situación contraria, es decir, con escasas exigencias (Karasek, 1979; Levi, 2001). Las condiciones de trabajo referentes a este estresor tienen que ver con la presión de clientes, las averías en equipos, la mala planificación de las tareas y del tiempo de realizarlas, la falta de personal, los materiales, herramientas y procesos de trabajo inapropiados, e incluso con la estructura salarial, por ejemplo, si para cobrar un mayor salario se trabaja más (Moncada, Llorens et al., 2014). Las exigencias psicológicas aparte de cuantitativas también tienen una dimensión cualitativa, que es la que afecta a los sentimientos, sobre todo cuando se necesita empatizar con las situaciones de otras personas, como pacientes, clientes o alumnos, tratando de no involucrarse y de no confundir los sentimientos de los otros con los sentimientos propios (exigencias psicológicas emocionales). También tiene que ver con las exigencias de esconder las emociones de cara al público, esto es, por requerimientos del puesto, se realiza el esfuerzo de aparentar una falsa emoción para mostrar un comportamiento neutral, con independencia del comportamiento de clientes, proveedores, compañeros de trabajo, superiores u otros usuarios ajenos o no a la empresa. Por un deseo de complacer o servir se suprime la emoción negativa (HaSPA, 2012; Moncada, Llorens et al., 2014; C. Nogareda, Gracia, Martínez y Salanova, 2007). Las exigencias cualitativas están relacionadas con la naturaleza de las tareas, por ese motivo no pueden eliminarse, así la manera de minimizar este estresor pasaría por reducir el tiempo de exposición y desarrollar habilidades para su afrontamiento (Moncada, Llorens et al., 2014).
- II. Conflicto trabajo-familia (doble presencia) (2 ítems, $\alpha=.73$). Esta dimensión hace referencia a la necesidad de atender las tareas laborales a la vez que a las domésticas y familiares, lo que supone una duplicidad de labores (doble presencia). Este doble trabajo puede constituir un riesgo para la salud, al igual que el conflicto de tiempo para intentar sincronizar o simultanear las tareas domésticas con la jornada laboral (Sánchez-Urán, 2004). El trabajo cotidiano doméstico-familiar socialmente es adjudicado al género femenino (Carrasquer, 2002), ese es el motivo de catalogar este factor de riesgo como específico de las mujeres (M. Rivas, 2008). Debido a esto, la gran parte de la literatura se enfoca en las trabajadoras, aunque algunos estudios, los menos, han relacionado la doble presencia también con hombres, como jóvenes con cargos de directivos (Weinberg,

Cooper y Weinberg, 1999). Por otra parte, la doble presencia impide a las mujeres dedicarse en pleno a su puesto de trabajo, generando buena parte de la desigualdad laboral entre hombres y mujeres, al limitar el acceso del sector femenino a determinados puestos de responsabilidad como los de dirección (Moreno y Garrosa, 2013), la inadecuada conciliación de la vida laboral y personal-familiar es un factor de riesgo para la salud.

Algunas de las condiciones laborales que pueden dar a origen a este estresor son las exigencias cuantitativas, la irregularidad, modificación o alargamiento de la jornada laboral, la falta de control sobre el horario de trabajo; flexibilización precaria de la jornada, los horarios incompatibles con la vida social o con el cuidado de las personas del entorno laboral, como por ejemplo el trabajo a turnos (Albertsen, Rafnsdóttir, Grimsó, Tómasson y Kauppinen, 2008; Moncada, Llorens et al., 2014).

- III. Control sobre el trabajo (Subdimensiones: influencia, posibilidades de desarrollo, sentido del trabajo) (6 ítems, $\alpha=.83$). Según Karasek (1979), el control sobre el trabajo está integrado por las dimensiones de influencia sobre el trabajo y el desarrollo de habilidades. Contar con autonomía sobre el contenido y las condiciones del trabajo es la base del trabajo activo y el aprendizaje, sin dejar de lado la importancia del sentido del trabajo (Moncada, Llorens et al., 2014).

La influencia tiene que ver con el margen de decisión y autonomía que posee el trabajador sobre el contenido y las condiciones laborales (orden, cantidad, tareas, métodos). Se refiere a *qué* tareas y *al cómo* desarrollarlas. En ese sentido está relacionado con la posibilidad que tiene el trabajador de tomar decisiones sobre los elementos más fundamentales de su trabajo. Unas altas demandas sumadas a una escasa capacidad de influencia en el trabajo, podría originar tensión psicológica y consecuencias negativas para la salud (De Lange et al., 2003; F. Jones y Fletcher, 2003; Karasek y Theorell, 1990). Cuando el trabajo constituye una oportunidad para desarrollar las habilidades y conocimientos del trabajador y/o adquirir otros nuevos, hablamos de las posibilidades de desarrollo, tiene relación con la influencia y con el diseño y los métodos del contenido de trabajo. Cualquier actividad laboral debería propiciar la adquisición de destrezas suficientes para poder realizar con éxito las tareas, así como permitir aplicar y mejorar los conocimientos y experiencias con los que ya cuenta el trabajador. Esta subdimensión está vinculada a la complejidad y variedad de las tareas, así, un trabajo rutinario o repetitivo sería fuente de exposiciones nocivas (HaSPA, 2012; Moncada, Llorens et al., 2014).

El sentido del trabajo es la tercera dimensión que compone el factor de control sobre el trabajo. Significa que aparte de la mera obtención de ingresos, el trabajo tiene sentido si puede relacionarse con otros valores o fines, como aprendizaje, utilidad o importancia social. Ser capaz de ver que las tareas que se realizan contribuyen con el servicio o producto final, y a su vez son compatibles con los propios valores, da significado a la actividad laboral. El contenido del trabajo y el significado de las tareas son las fuentes directas del sentido del trabajo, el cual supone un factor de protección frente a otras posibles exigencias laborales (Moncada, Llorens et al., 2014). Por el contrario, una escasa valoración del trabajo es una potencial fuente de estrés (Kasl, 1992).

- IV. Apoyo social y calidad de liderazgo (Subdimensiones: previsibilidad, claridad de rol, calidad de liderazgo, conflicto de rol) (8 ítems $\alpha=.87$). El apoyo social implica disponer de la ayuda (instrumental y/o emocional) necesaria en el trabajo por parte de compañeros, subordinados y superiores, tanto en términos de calidad como de cantidad, también supone recibirla en el momento adecuado (HaSPA, 2012; Sauter, Murphy y Hurrell,

1992). El apoyo social es de los elementos más deseados en el trabajo por su efecto amortiguador de las consecuencias del estrés (Karasek, Schwartz y Theorell, 1982; Peiró y Salvador, 1993; Viswesvaran, Sánchez y Fisher, 1999). Las tres subdimensiones de la versión corta que componen el apoyo social son la previsibilidad, la claridad de rol y el conflicto de rol.

La previsibilidad significa poseer la información suficiente en tiempo y forma para poder realizar correctamente el trabajo y adaptarse a los cambios que afecten a la esfera laboral, como nuevas tecnologías o nuevos métodos o procedimientos. Este factor se relaciona con la falta de formación, información y apoyo a los cambios (Moncada, Llorens et al., 2014).

La claridad de rol tiene que ver con la definición del puesto de trabajo, es decir, el trabajador conoce las tareas a realizar, los recursos, los objetivos, el alcance de sus responsabilidades en el trabajo y lo que se espera de él. La falta de claridad en este aspecto supone ambigüedad de rol, que constituye una amenaza para las exigencias de adaptación del puesto (Jex, 2001; Moncada, Llorens et al., 2014). La falta de claridad de rol suele ocurrir cuanto un puesto de trabajo o sus tareas están mal definidas desde su origen, también en situaciones de reestructuración empresarial o cuando se añaden equipos o desarrollos de productos nuevos (Ivancevich y Matteson, 1980). Otros motivos que convierten la ambigüedad de rol en un estresor, son la falta de autonomía y la comunicación deficiente entre trabajadores y superiores (Jex, 2001).

El conflicto de rol se manifiesta cuando existen demandas contradictorias que el trabajador no desea cumplir, o cuando debe desempeñar funciones que contradicen sus valores éticos o profesionales, entonces se pone en duda el realizar las tareas o no. Esta falta de decisión de una forma clara y rápida supone un conflicto que actúa como un importante estresor que disminuye la satisfacción laboral (Cuenca, 2002; Kahn, et al., 1964) y aumenta el riesgo de padecer enfermedades (Ivancevich y Matteson, 1980).

La calidad de liderazgo es el papel que desempeña la dirección y los mandos intermedios en cuanto a la política empresarial de gestión de la mano de obra. Este factor viene determinado por la cultura de la organización y por las características de los líderes, en ese sentido, la calidad de dirección es la responsable de propiciar la motivación, el bienestar y el crecimiento personal de los trabajadores, así como la aceptación y el seguimiento por parte de éstos (Bass, 1990). Un liderazgo inadecuado dependerá a su vez de varios factores como las características de la tarea, o las propias del trabajador, pero de modo general, entre algunas de las consecuencias que entraña se encontraría la baja autoestima, insatisfacción, incertidumbre, escaso de desarrollo de las habilidades personales (Cooper y Smith, 2001).

- V. Compensaciones del trabajo (Subdimensiones: inseguridad sobre el empleo, inseguridad sobre las condiciones de trabajo (4 ítems $\alpha=.73$). Esta dimensión está basada en el modelo esfuerzo-recompensa propuesto por Siegrist (1996a, 1998, descrito en el epígrafe 2.3.1.2), que sostiene que a largo plazo, si el trabajador tiene que realizar esfuerzos elevados unido a bajas compensaciones, este desequilibrio puede actuar como estresor con el consiguiente riesgo para su salud (HaSPA, 2012).

En la versión corta del CoPsoQ-ISTAS 21 se incluyen dos subdimensiones, la primera de ellas es la inseguridad sobre el empleo, se trata de la preocupación por el futuro laboral y está muy relacionada con la estabilidad en el empleo y las posibilidades de encontrarlo (Moncada, Llorens et al., 2014). Entre las condiciones laborales que pueden ser fuente de

inseguridad en el empleo se encuentra la temporalidad, la reestructuración de la empresa o los despidos. Estas cuestiones son percibidas por el trabajador como amenazas y en consecuencia operan como estresores (Ivancevich, 2001).

La segunda de las subdimensiones es la inseguridad sobre las condiciones de trabajo. Abarca la inquietud sobre el empeoramiento de las condiciones laborales y está relacionado con cambios en las condiciones sustanciales del trabajo como la movilidad funcional y geográfica, los cambios de la jornada y horario de trabajo, el salario y la forma de pago o la carrera profesional (Moncada, Llorens et al., 2014).

- VI. Capital social (Subdimensiones: Justicia, confianza vertical) (4 ítems, $\alpha=.88$). Bourdieu (1986), Coleman (1988) y Putnam (1993) son considerados los pioneros de este concepto, que puede definirse como “el conjunto de recursos colectivos que permiten a sus miembros solucionar entre todos, la tarea clave de ésta” (Moncada, Llorens et al., 2014, p. 36). Debido a los distintos niveles de jerarquía que coexisten entre los distintos miembros de la organización, la colaboración entre todos no es posible sin la confianza que se genera a raíz de la percepción de justicia organizacional (Olesen, Thoft, Hasle y Kristensen, 2008).

La justicia es la medida en que los trabajadores son tratados con equidad. Una de las fuentes de falta de justicia se origina en las prácticas de la gestión laboral, como la distribución de las tareas, la solución de conflictos o la promoción. La percepción de falta de justicia se relaciona principalmente con perturbaciones de tipo psiquiátrico y depresión entre otras dolencias (Kivimäki, Elovainio, Vahtera, Virtanen y Stansfeld, 2003; Ylipaavalniemi, Kivimäki, Elovainio, Virtanen, Keltikangas-Järvinen y Vahtera, 2005).

Por último, la confianza vertical se refiere a la confianza que se deposita en la dirección y trabajadores de que actuarán de manera adecuada y competente. Se refiere a la seguridad de que las personas con poder no se aprovecharán de las vulnerables. La falta de confianza vertical está relacionada la experiencia de justicia organizacional, trato equitativo y transparencia (Acosta, Torrente, Llorens y Salanova, 2015; Moncada, Llorens et al., 2014).

5.3.5.2. Escala de afecto positivo y negativo (PANAS)

La escala *Positive and Negative Affect Scale* (PANAS), desarrollada originalmente por Watson et al. (1988), tiene una estructura bidimensional que evalúa el AN y el AP. Una de las primeras adaptaciones a la población española fue realizada por Sandín et al. (1999) y años más tarde por Godoy-Izquierdo, Martínez y Godoy (2008). La escala está formada por dos dimensiones independientes y por lo tanto no están correlacionadas. Está compuesta por 20 ítems descriptores de sentimientos y emociones, 10 de ellos se utilizan para medir AN y los otros 10 para AP. En el presente trabajo sólo se utiliza la dimensión de AN, que se refiere a la experimentación de emociones desagradables (ej., enojado, tenso, nervioso) (10 ítems, $\alpha=.81$).

Los informantes responden según el grado en el que, habitualmente, perciban las distintas emociones mediante escala tipo Likert de 5 puntos con los siguientes calificadores: nada (1), poco (2), algo (3), bastante (4) y mucho (5). Como las puntuaciones de cada uno de los grupos se suman separadamente, sólo se van a tener en cuenta los índices relativos al AN.

5.3.5.3. Ítems de control de la deseabilidad social

Como se ha descrito en el marco teórico, multitud de medidas de autoinforme referidas a personalidad y salud son susceptibles a la DS (Enríquez y Domínguez, 2010; González e Iruarrizaga, 2005), afectando también al ámbito de la psicología del trabajo y de las organizaciones (Salgado, 2005).

Con objeto de poder identificar y medir esta tendencia de respuesta, se han incluido 4 ítems a modo de marcadores. Son los mismos reactivos utilizados en el método de Ferrando (2005), pues ese mismo procedimiento es el que se pretende utilizar para el estudio de este sesgo de respuesta. Los ítems se describen en la siguiente tabla.

Tabla 60: Marcadores de deseabilidad social.
Fuente: Elaboración propia

MARCADORES DS	POSICIÓN EN EL CUESTIONARIO
Alguna vez he dicho algo malo de otra persona	ítem 15
A veces dejo para mañana lo que debo hacer hoy	ítem 16
Alguna vez me he aprovechado de alguien.	ítem 17
Alguna vez he cogido alguna cosa que no era mía	ítem 18

Al igual que el resto de reactivos que componen el cuestionario, el formato de respuesta para los marcadores DS ha sido el de la escala Likert de 5 puntos. Se colocaron intercalados en la escala CoPsoQ para reducir las respuestas por asociaciones causales y minimizar así el sesgo de varianza de método común (Podsakoff et al., 2003).

5.3.5.4. Fiabilidad de las escalas de medida empleadas

Se ha utilizado el índice alfa de Cronbach para examinar la consistencia interna de las escalas empleadas en la investigación (ver tabla 61). Según la categorización de George y Mallery (2007) el alpha de Cronbach se interpretará de acuerdo a la siguiente escala: >.90 excelente, >.80 optimo, >.70 aceptable, >.60 cuestionable, >.50 pobre y <.50 inaceptable.

Aunque se considera como aceptable tomar valores a partir de 0.7, se ha demostrado aceptables valores inferiores en investigaciones exploratorias (Hair, Anderson, Tatham y Black, 2007; Huh, Delorme y Reid, 2006), otros autores consideran suficientes valores de .50 y .60 en las primeras fases de investigación (Nunnally y Bernstein, 1995). En orden a estos criterios, y tal como se observa en la mencionada tabla 61, todas las escalas de esta investigación tienen valores superiores a .60, a excepción de las exigencias emocionales cuyo valor de alfa Cronbach es de .49.

Tabla 61: Fiabilidad de las escalas de medida
Fuente: Elaboración propia

Instrumentos de medida	Dimensiones	Nº Ítems	α
ISTAS 21 (CoPSoQ) v2		30	.90
	Exigencias psicológicas en el trabajo:	6	.77
	Exigencias cuantitativas	2	.75
	Ritmo de trabajo	2	.72
	Exigencias emocionales	2	.49
	Conflicto trabajo familia	2	.73
	Control sobre el trabajo	6	.83
	Influencia	2	.70
	Posibilidades de desarrollo	2	.81
	Sentido del trabajo	2	.88
	Apoyo social y calidad de liderazgo	8	.87
	Previsibilidad	2	.79
	Claridad de rol	2	.84
	Conflicto de rol	2	.79
	Calidad de liderazgo	2	.91
	Compensaciones del trabajo	4	.73
	Inseguridad sobre las condiciones de trabajo	2	.66
	Inseguridad sobre el empleo	2	.71
	Capital social	4	.88
	Justicia	2	.87
	Confianza vertical	2	.80
PANAS	Afecto negativo	10	.81
Deseabilidad Social	Marcadores DS	4	.63

5.3.6. Análisis de datos

Para validar o refutar las hipótesis planteadas se ha utilizado como soporte los siguientes paquetes estadísticos Statistical Package for Social Sciences (IBM SPSS) versión 24 para Windows; Matlab; Excel; Factor; G*Power y AMOS.

En primer lugar se comprobó si existía sesgo de método común mediante el test de factor único de Harman y seguidamente se calcularon los estadísticos descriptivos para poder categorizar demográficamente la muestra. Los siguientes análisis se enfocaron para determinar, (1) si la DS tiene algún efecto en las medidas de autoinforme de los riesgos psicosociales en el trabajo y (2) el funcionamiento diferencial del ítem para comprobar la existencia de posibles sesgos producidos por el rasgo de AN.

Para poder realizar los cálculos referentes a la identificación del sesgo de deseabilidad social en el cuestionario CoPsoQ-ISTAS 21, se hizo necesario proponer un nuevo modelo factorial debido al desconocimiento de los datos de la estructura factorial de origen. Para comprobar la idoneidad de la propuesta factorial se dividió la muestra total en dos submuestras aleatorias. A la primera submuestra se le aplicó el análisis paralelo de Horn (1965), con objeto de conseguir una primera aproximación de la reducción factorial más idónea. Con esos

resultados se estableció un Análisis factorial exploratorio (AFE). En la segunda de las muestras se realizó un Análisis factorial confirmatorio (AFC) con el que se determinó el número de factores óptimos que mejor explicaban la varianza compartida por los ítems.

5.3.6.1. Sesgo de varianza común del método

Los cuestionarios de autoinforme pueden ser una potencial fuente de sesgo debido al método, en el intento de minimizar esa varianza se ha seguido a Podsakoff et al. (2003) llevando a cabo técnicas estadísticas y no estadísticas. Las no estadísticas han sido: (i) favorecer la separación psicológica de variables dependientes e independientes, utilizando para ello encabezados que diferencian de forma clara las distintas partes del cuestionario (Podsakoff, MacKenzie, y Podsakoff, 2012); y (ii) garantizar la confidencialidad y proteger el anonimato, pues de este modo se reduce el sesgo, incluso de respuestas en temas sensibles (Konrad y Linnehan, 1995).

En cuanto a los procedimientos estadísticos se ha utilizado el test de factor único de Harman (Meade et al., 2007; Konrad y Linnehan, 1995; Pavlou et al., 2007; Podsakoff y Organ, 1986) para detectar la existencia (o no) de varianza común de método. Este test consiste en la aplicación de un análisis factorial a todas las variables del cuestionario mediante el método de extracción de factorización por ejes principales (Ylitalo, 2009). El procedimiento implica introducir las variables de todas las escalas de interés (CoPsoQ-ISTAS 21 II, PANAS, DS) y forzarlas para la extracción de un único factor. De este modo se comprueba si existe un factor único de varianza debido, por ejemplo, a la utilización del mismo sistema de medida (escala Likert de 5 puntos) para evaluar distintas variables (Littlewood, 2006).

5.3.6.2. Efecto de la Deseabilidad Social en la escala CoPsoQ-ISTAS 21 v.2

El procedimiento de análisis que se ha seleccionado para controlar el efecto de la DS, se basa en el método descrito por Ferrando (2005) y Ferrando et al. (2009), que es una extensión del método de Neil y Jackson (1970) basado en técnicas de análisis factorial, que ha sido utilizado en distintos estudios (Morales-Vives, Camps y Lorenzo-Seva, 2013; Ruiz-Pamies, 2014; Ruiz-Pamies, Lorenzo-Seva, Morales-Vives y Vigil-Colet, 2014). Lo que distingue este proceso de otros, es que consigue extraer las saturaciones de DS en los ítems de contenido, consiguiendo puntuaciones individuales limpias de sesgo. De este modo consigue salvar las limitaciones de otros métodos de control, como los que utilizan escalas específicas de detección de DS donde es fácil eliminar información de interés al corregir el sesgo, pues a menudo se descartan las respuestas de los sujetos con altos niveles en la variable, corriendo el riesgo de estar eliminando puntuaciones verdaderas (Morales, 2006).

Los análisis utilizados por Ferrando (2005) y Ferrando et al. (2009) se basan en dos fases contiguas, donde inicialmente se extrae el efecto DS y a continuación se aplica un análisis factorial sobre las puntuaciones libres de dicho sesgo, y así se consigue extraer los factores de contenido a evaluar; en nuestro caso el nivel de exposiciones a los factores de riesgo psicosocial.

Para poder llevar a cabo este método de control, dada la ausencia de datos sobre la estructura factorial de la versión corta de la escala CoPsoQ, se han realizado análisis factoriales exploratorios y confirmatorios, con objeto de conocer la dimensionalidad de la misma. Para

conseguir evidencias acerca de la capacidad predictiva del modelo factorial propuesta (validez cruzada), se ha dividido la muestra original en dos submuestras al azar:

Submuestra 1= 60% aprox. de los individuos para realizar un AFE ($n_1=325$).

Submuestra 2= 40% aprox. de los individuos para realizar un AFC ($n_2=238$).

Una vez extraídos los factores de la escala se procede a aplicar el método de Ferrando, asumiendo que no existe correlación entre los factores de contenido y el factor DS (Couch y Keniston, 1960; Gough y Heilbrun, 1983; Messick, 1966; Nunnally y Bernstein, 1995). La primera etapa del método se basa en una matriz inicial de correlaciones inter-ítem con la influencia de DS, cuya modelo estructural se representa en la siguiente ecuación:

$$X_{ij} = \lambda_{j1} \theta_{i1} + \lambda_{j2} \theta_{i2} + \varepsilon_{ij}$$

Donde, q_1 = factor de contenido; q_2 = factor de Deseabilidad Social; l = cargas factoriales; y e = residuos

La ecuación muestra una solución factorial de un único factor de contenido, donde cada reactivo está compuesto por la suma de (1) las cargas factoriales del factor de contenido, (2) la suma de las cargas del factor DS, y (3) los residuos. Los 4 ítems utilizados de marcadores de DS permitirán identificar el factor relacionado con la DS dentro de la matriz inter-ítem. Una vez obtenida la matriz de correlaciones inter-marcadores con el programa MATLAB, se efectúa un análisis factorial semiestricto para obtener los valores de las cargas de cada marcador en el factor DS, con dichos valores se calculan los pesos factoriales de los ítems de contenido en el factor DS, utilizando para ello la *Técnica Variable Instrumental* (Hägglund, 1982). Por último, se descarta la varianza explicada por el factor DS de la matriz de correlaciones inter-ítem, consiguiendo de este modo una matriz limpia del sesgo de DS (matriz residual).

En la segunda etapa se parte de la matriz residual y se realiza un análisis factorial exploratorio utilizando el programa FACTOR (Lorenzo-Seva y Ferrando, 2013) a partir del cual se consiguen las estimaciones de la proporción de la varianza común y la varianza explicada para cada factor de contenido mediante el método de extracción de factores de *Minium Rank Factor Analysis* (MRFA, Ten Berge y Kiers, 1991). Sólo puede obtenerse mediante análisis factorial exploratorio o semi-confirmatorio, pues el AFC no distingue entre la varianza común (comunalidad) y la varianza error (unicidad) (Lloret-Segura et al., 2014; Ferrando y Anguiano-Carrasco, 2010). La estructura factorial obtenida se rota mediante la técnica de rotación oblicua Promin (Lorenzo-Seva, 1999) para permitir la correlación entre factores. Efectuado este análisis se distingue una solución donde cada reactivo mostrará una carga en el factor de contenido y otra, preferiblemente secundaria, en el factor ortogonal de DS, obteniendo así puntuaciones limpias del sesgo DS.

5.3.6.3. Funcionamiento Diferencial de Ítem.

El objetivo de utilizar este análisis es el de comprobar la posible existencia de un funcionamiento diferencial de los ítems en función de la variable de AN, en la versión corta del CoPsoQ-ISTAS 21 V.2. El principal requisito que debe cumplir cualquier instrumento de medida es la objetividad, por consiguiente, no puede estar afectado por las características del objeto o sujeto que va a evaluar. Si fuese así, su validez estaría en entredicho. En ese sentido,

uno de los procedimientos utilizados para conocer la validez de las pruebas aplicadas es explorar la existencia de funcionamiento diferencial del ítem *Differential Item Functioning* (DIF), vocablo acuñado por Holland y Thayer (1988).

El DIF se refiere a la posibilidad de responder a un ítem de modo distinto dependiendo de la pertenencia o no a un subgrupo de una población dada (ej., género, clase social, etnia, sexo), partiendo de sujetos comparables; es decir, sujetos diferentes, pero con el mismo nivel en el atributo o rasgo que mide el instrumento. Si esto ocurre, el test no estaría garantizando los mismos resultados para todos los sujetos, ya que estaría funcionando distintamente a favor o en contra de un determinado grupo (Fidalgo, Mellenbergh y Muñiz 1998; Fidalgo y Paz, 1995). Esto es, las diferencias observadas no se corresponden con diferencias equivalentes en la variable estudiada o en el nivel de verdad. Por ejemplo, respecto al DIF de género, se ha encontrado que las mujeres responden de forma más variada e intensa a reactivos que impliquen sentimientos que los hombres, sin que esas diferencias se correspondan con las verdaderas percepciones sobre la experiencia emocional (Fischer, 2000; Pennebaker, Rime y Blankenship, 1996; Thissen, Steinberg y Gerrard, 1986).

Habitualmente, los tipos de DIF se agrupan en dos categorías, DIF uniforme y DIF no uniforme (Mellenbergh, 1982, 1995). Se considera uniforme cuando la probabilidad de contestar de forma correcta a un ítem es mayor para un grupo que para otro, en cada uno de los niveles de la variable. Por el contrario, el DIF se dice que no es uniforme cuando la posibilidad de responder correctamente al ítem entre dos grupos no es igual en todos los niveles de la variable (Gómez-Benito, Hidalgo y Ferré, 2005).

Pueden ser comparados varios grupos de sujetos, aunque normalmente se centran en dos, por una parte, el grupo de referencia (R) referido al grupo mayoritario o aventajado, y por otra parte, el grupo focal (F) que sirve de criterio, sería el grupo minoritario que se encuentra en desventaja comparándolo con el grupo de referencia.

Aunque en la bibliografía revisada, las investigaciones sobre DIF se refieren fundamentalmente a variables sociodemográficas –concretamente, el estudio más extendido versa sobre diferencias de género-, en este trabajo se pretende averiguar la existencia o no de funcionamiento diferencial en el ítem en sujetos con puntuaciones altas en AN en contraposición a los que puntúan bajo. Es decir, se analizará si sujetos bajos y altos en AN, igualados en el nivel de exposición a los factores psicosociales, responden con diferente frecuencia a las alternativas de respuesta de la mayoría de los ítems del CoPsoQ-ISTAS21.

En cuanto a los procedimientos para la detección del DIF existen muy diversas técnicas como estadísticos χ^2 tradicionales, extensiones de la regresión logística, generalizaciones del estadístico Mantel-Haenszel y otras (Gómez-Benito e Hidalgo, 1997; Gómez-Benito, et al. 2005). En el presente estudio se utilizará el método basado en comparación de modelos, propuesta por T.R. Miller y Spary (1993) de análisis de discriminante logístico (ADL), por ser los ítems del cuestionario de respuesta ordenada tipo Likert, es decir, politómicos. Este procedimiento se ha demostrado eficaz para detectar tanto DIF uniforme como no uniforme (Hidalgo y Gómez-Benito, 2000; Hidalgo y López-Pina, 2004; Zumbo, 1999). Está basado en un análisis de regresión logística donde se obtienen, para cada ítem, tres modelos sucesivos de regresión, tomando como variable dependiente el grupo de pertenencia (g) (en el presente estudio, AN alto o AN bajo), y como variables predictoras, (i) la respuesta al ítem, (ii) la

puntuación total en la escala analizada (CoPsoQ) (θ) y (iii) la interacción entre los dos factores ($\theta * g$).

Para el estudio que nos ocupa la variable predictora será el AN y para su análisis DIF se seguirán los criterios teóricos y metodológicos expuestos por Hidalgo, Gómez y Padilla (2005) y utilizados en el trabajo de Guillén-Riquelme y Buela-Casal (2011). En las líneas que siguen se describe más detalladamente la estrategia a seguir. Se basa en comparar las razones de verosimilitud de los de modelos anidados, ajustándolos en distintas fases, modelo 1 (M1), modelo 2 (M2) y modelo 3 (M3), comprobando la significación del efecto en cada modelo. La ecuación de regresión sería la siguiente:

$$Y = \beta_0 + \beta_1\theta + \beta_2g + \beta_3\theta * g$$

Procediendo del modo que sigue:

- Paso 1: Se ajusta el M1 añadiendo a la ecuación la puntuación total de cada sujeto en el test (θ), que sería la variable que indica el nivel de habilidad.
- Paso 2: Se configura el M2 incorporando a la ecuación la variable de pertenencia al grupo (g). Si este segundo modelo respecto al primero fuese estadísticamente significativo se demostraría la presencia de DIF uniforme.
- Paso 3: Para el M3 se agrega la interacción entre la puntuación total en el test y la variable dependiente, que es la variable de grupo ($\theta * g$). Si el resultado de comparar este modelo con el anterior M2 resultase estadísticamente significativa, nos encontraríamos ante DIF no uniforme.

Esta prueba conjunta de detección de DIF uniforme y no uniforme ha presentado mejores resultados en la detección correcta de ítems afectados con DIF y en controlar el error del tipo I (falsas detecciones) (Hidalgo et al., 2005). Estadísticamente cada modelo proporciona el valor de Chi-cuadrado (χ^2) y para considerar que un ítem entre los grupos (AN) está afectado por DIF tiene que cumplirse que el valor de χ^2 sea significativamente mayor al valor χ^2 del primer paso. Se considera que las diferencias serán significativas cuando el p-value para χ^2 es menor o igual a 0.01, es decir, el incremento de χ^2 en el M2 tendría que resultar superior a 9.21 ($p = 0.01$) (Gelin y Zumbo, 2003).

La aplicación del método de regresión logística de T.R Miller y Spary (1993), además de un test de significación, permite el cálculo del efecto de DIF, basado en las diferencias R^2 de Nagelkerke entre dos modelos, concretamente M2 y M1 (Thomas y Zumbo, 1996). A parte de la significación, es necesario, para considerar la existencia de DIF, que χ^2 sea de magnitud suficiente, a este respecto, siguiendo las propuestas de Jodoin y Gierl (2001) se tomarán en cuenta los siguientes parámetros descritos en la tabla 62.

Tabla 62: Criterios de magnitud de DIF
Fuente: Adaptado de Jodoin y Gierl (2001)

ΔR^2 Nagelkerke	Magnitud
$R^2 < 0.035$	Inapreciable
$0.035 < R^2 < 0.070$	Moderado
$R^2 > 0.070$	Elevado

R^2 significa la proporción de variación de las respuestas al ítem explicada por la ecuación de regresión, por lo tanto, siguiendo los criterios descritos en la tabla anterior, aparte de la significación del χ^2 el ítem mostrará DIF elevado si la diferencia de R^2 entre M2 y M1 respectivamente es superior a 0.07 y será considerado DIF moderado si la diferencia se encuentra entre 0.035 y 0.070.

Como último paso se efectuará la depuración de los ítems con DIF moderado o alto por etapas, para ello se eliminarán las puntuaciones de los ítems afectados del total del cuestionario y una vez corregido el test se volvería a repetir la ecuación de regresión con objeto de averiguar la existencia de nuevos ítems con DIF y en ese caso se volvería a purificar. La purificación de ítems en regresión logística disminuye el error tipo del tipo I a la vez que aumenta el poder de detección (B. French y Maller, 2007). Para averiguar si el DIF se presenta en todas las escalas se utiliza la técnica propuesta por Mellenbergh (1995) de categorías acumuladas. El análisis consiste en varias dicotomizaciones de las categorías de respuesta, repitiendo los análisis de DIF en cada una de las agrupaciones.

5.4. RESULTADOS

5.4.1.Exploración de la varianza común de método

El porcentaje de varianza explicada por el factor fue de 22.93, muy inferior al nivel de tolerancia indicado por Ylitalo (2009) para datos con menos de 500 observaciones (63.30%). Otra variante muy utilizada dentro del test de Harman es mediante el método de extracción por componentes principales. Aplicado a las variables del cuestionario se obtuvieron 11 factores con autovalores superiores a 1 que explicaban el 62.63% de la varianza total. Al haberse identificado la presencia de más de un factor y dado que el primer factor no explicaba la mayoría de la varianza se puede presuponer que las respuestas no han sido facilitadas de forma sistemática y por lo tanto el sesgo de la varianza del método común parece no ser un contaminante probable que afecte esta investigación (Podsakoff y Organ, 1986).

5.4.2.Análisis descriptivos

Con el propósito de obtener una visión general de la calidad de los ítems de la escala CoPsoQ y de los marcadores de DS, en la tabla 63 se muestran los principales estadísticos descriptivos de la submuestra 1 ($n_1=325$), que incluyen la media, la desviación típica, la asimetría y la curtosis.

Tabla 63: Estadísticos descriptivos de los ítems del CoPsoQ-ISTAS21 y de DS ($n_1=325$)

Fuente: Elaboración propia

Ítem	\bar{x}	σ	As	K
DS1 +	1.70	0.75	1.17	0.78
DS2 +	2.08	0.65	0.25	-0.63
DS3 +	1.36	0.35	1.61	2.39
DS4 +	2.68	0.81	-0.10	-0.70
CoPsoQ1 +	2.25	1.09	-0.32	-0.38
CoPsoQ2 -	1.74	1.22	0.25	-0.52
CoPsoQ3 +	1.75	1.11	-0.18	-0.82
CoPsoQ4 +	1.855	1.37	0.02	-0.86
CoPsoQ5 +	1.65	1.35	0.21	-0.86
CoPsoQ6 +	2.39	1.10	-0.20	-0.48
CoPsoQ7 -	2.42	1.29	-0.34	-0.80
CoPsoQ8 -	2.83	1.04	-0.82	0.25
CoPsoQ9 +	2.32	1.50	-0.27	-0.84
CoPsoQ10 +	2.44	1.34	-0.31	-0.82
CoPsoQ11 -	2.74	1.32	-0.57	-0.66
CoPsoQ12 -	2.99	1.05	-0.97	0.44
CoPsoQ13 -	3.08	0.94	-0.89	0.02
CoPsoQ14 -	3.06	1.06	-1.05	0.45
CoPsoQ15 -	2.87	0.97	-0.55	-0.52
CoPsoQ16 -	2.72	1.32	-0.60	-0.60
CoPsoQ17 +	1.70	1.50	0.36	-0.86
CoPsoQ18 +	2.19	1.40	-0.04	-0.92
CoPsoQ19 -	1.67	1.45	0.45	-0.81
CoPsoQ20 -	2.11	1.29	0.09	-0.96
CoPsoQ21 +	0.81	1.48	1.35	0.58
CoPsoQ22 +	1.46	2.23	0.52	-1.22
CoPsoQ23 +	1.57	2.43	0.45	-1.32
CoPsoQ24 +	2.18	2.15	-0.12	-1.40
CoPsoQ25 -	2.61	1.03	-0.49	-0.43
CoPsoQ26 -	2.39	1.22	-0.33	-0.64
CoPsoQ27 -	2.01	1.28	0.14	-0.81
CoPsoQ28 -	1.88	1.32	0.12	-0.84
CoPsoQ29 -	1.96	1.47	-0.07	-0.98
CoPsoQ30 -	1.98	1.56	-0.02	-1.04

 \bar{x} = Media; σ = Desviación típica; As= Asimetría, K= curtosis

Respecto al análisis de normalidad, los valores de asimetría univariada estuvieron comprendidos entre -1.05 y 1.61, y los valores de curtosis univariada entre -1.40 y 2.39, el promedio de asimetría es de 0.28 y el de curtosis -0.60. Estos datos cumplen las criterios de normalidad establecidos por diversos autores (Chou y Bentler, 1995; Curran, West y Finch, 1996; West, Finch y Curran (1995), ya que los datos para valores de asimetría univariada no son superiores a ± 3 ni los de curtosis mayores de ± 7 , también se cumplen otros criterios de asimetría más estrictos, como los comprendidos entre el rango $[-2, 2]$ (Bandalos y Finley, 2010), o los que contemplan el promedio de la asimetría y curtosis expresadas en valor absoluto en

valores menores a uno (Muthen y Kaplan, 1992). Dado que (1) los ítems son politómicos con un mínimo de cinco opciones de respuesta, y (2) la distribución univariada de los ítems no es asimétrica ni presenta exceso de curtosis, se calculó la matriz de correlaciones de Pearson entre los ítems del QoPsoQ-ISTAS21 y los marcadores de DS como por recomienda en Lloret-Segura et al. (2014).

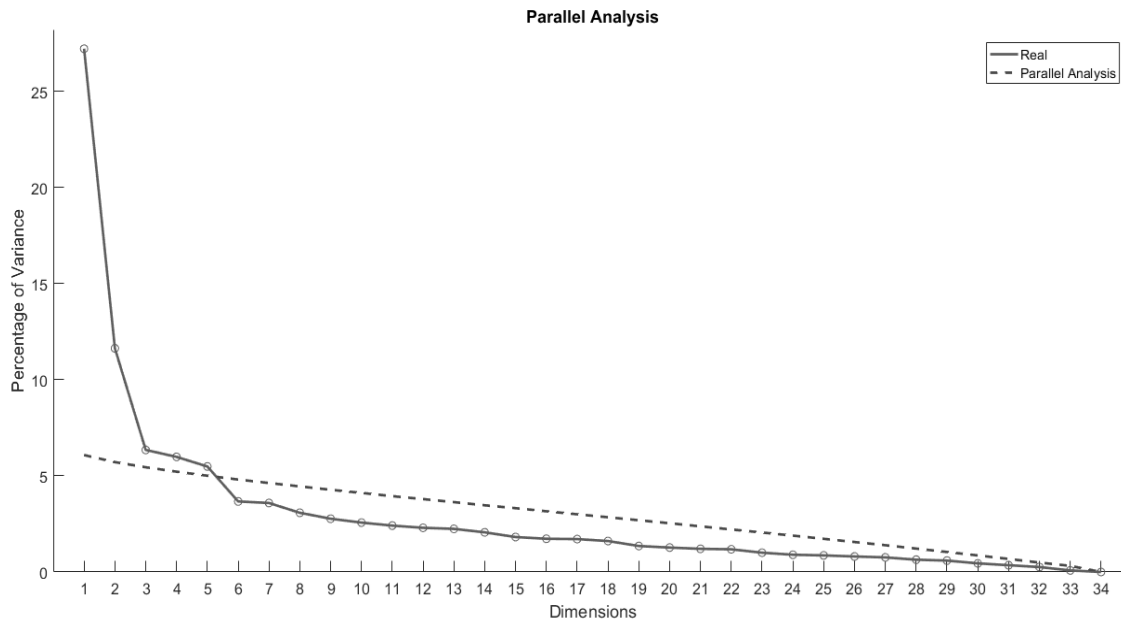
5.4.3. Depuración del sesgo de Deseabilidad social

Ante la ausencia de información previa sobre la estructura factorial de la versión reducida del CoPsoQ 21-ISTAS21 v.2 y dada la necesidad de conocer dichos datos para aplicar el método de Ferrando, se opta por realizar un análisis factorial a la matriz de correlaciones obtenida y así poder conocer el número de dimensiones subyacentes a los datos. Para realizar estos cálculos se ha seleccionado la primera submuestra al azar ($n_1=325$) que, aproximadamente, corresponde al 60% de la muestra original.

El primer paso fue determinar si la matriz cumple las condiciones para ser factorizada comprobando el grado de relación conjunta entre las variables, para ello se calculó el estadístico de Kaiser- Meyer-Olkin (KMO). El rango del KMO oscila entre 0 y 1, y cuanto más se aproxima a 1, más relación sustancial habrá entre las variables y, por consiguiente, se cumplirán los supuestos para aplicar el análisis factorial. Según Kaiser (1970) y Kaiser y Raice (1974) el valor sugerido para que la matriz a factorizar sea satisfactoria es $\geq .80$. El índice KMO de la matriz de correlaciones analizada arrojó un valor de .86, lo que implica que las respuestas no son casuales y que los resultados de la factorización serán estables.

Para determinar el número de dimensiones subyacentes a los datos se calculó el análisis paralelo (Horn, 1965) recomendado en Lloret-Segura et al. (2014). Para ello se empleó el programa FACTOR (Lorenzo-Seva y Ferrando, 2013), que compara cada autovalor con autovalores aleatorios con la finalidad de retener sólo aquellos factores que están relacionados con una cantidad de varianza mayor que la cantidad de varianza de los factores obtenidos al azar. De acuerdo con Lattin, Carroll y Green (2003), se sugirieron 5 dimensiones subyacentes a los datos. Estos resultados proponen una dimensión relacionada con el estilo de respuesta (DS) y cuatro dimensiones de contenido concernientes al CoPsoQ-ISTAS21. El gráfico de sedimentación puede verse en la figura 11, donde el punto de intersección entre la varianza generada al azar y la empírica muestra el número de factores a retener.

Figura 11. Resultados del análisis paralelo (n=325).
Fuente: Elaboración propia



De este modo, se llevó a cabo un AFE sobre la matriz de correlaciones inter-ítem y la matriz de correlaciones residual; esto es, después de eliminar la varianza debida a la DS una vez aplicado el procedimiento desarrollado por Ferrando et al. (2009). De este modo se retuvieron cuatro dimensiones de contenido empleando el método de análisis de mínimo factor de rango (MRFA, Ten Berge y Kiers, 1991). Esta técnica permite valorar la proporción de la varianza común explicada al distinguirla de la varianza total común. Con objeto de maximizar la simplicidad factorial se utilizó el criterio Promin de rotación oblicua (Lorenzo-Seva, 1999), este método tiende a conseguir soluciones lo más simples posibles incluso en modelos complejos, por ello y por su refinamiento, es una técnica recomendada sobre otras (Ferrando, Anguiano-Carrasco, 2010; Lloret-Segura, et al., 2014; Tous-Pallarès, Bonasa-Jiménez, Mayor-Sánchez y Espinoza-Díaz, 2011). Debido al análisis paralelo (Horn, 1965) efectuado con anterioridad, se probó la estructura hipotetizada reteniendo cuatro factores de contenido, que fueron etiquetados como sigue: *Control sobre el trabajo* (F1), *Exigencias psicológicas en el trabajo* (F2), *Compensaciones del trabajo* (F3) y *Apoyo social y calidad de liderazgo* (F4), y que se describen a continuación:

- El primer factor (F1) se ha denominado *Control sobre el trabajo* y está formado por los ítems 7, 8, 11, 12, 13, 14, 15 y 16, con una fiabilidad factorial de .91. De forma general, trata sobre el grado de autonomía que tiene el trabajador sobre el modo de realizar sus funciones. Siguiendo la versión original del CoPsoQII (Pejtersen, et al., 2010) engloba las subdimensiones de *influencia* (ítems 7 y 8), *posibilidades de desarrollo* (ítems 11 y 12) y *sentido del trabajo* (ítems 13 y 14). Como resultado del AFE también incluye la subdimensión de *la claridad de rol* (ítems 15 y 16). La claridad de rol está relacionada con el conocimiento de los aspectos fundamentales de la actividad laboral, como tareas, funciones, equipos, responsabilidades o margen de autonomía en el trabajo (Moncada,

Llorens et al., 2014), el desconocimiento de estos aspectos no facilitaría el control sobre el trabajo, por lo que su encuadre dentro de este primer factor parece razonable.

- El segundo factor (F2) se ha definido como *Exigencias psicológicas en el trabajo*, su fiabilidad factorial es de .87 y se compone de los ítems 1, 2, 3, 4, 5, 6, 9 y 10. Se refiere a las demandas laborales, tanto cuantitativas como emocionales. Concuerda con el marco teórico de partida que está compuesto por las subdimensiones de *exigencias cuantitativas* (ítems 1 y 2), exigencias emocionales (ítems 5 y 9) y ritmo de trabajo (ítems 6 y 10). A su vez, incluye los ítems 3 y 4 que se corresponden con la dimensión *conflicto trabajo-familia*. Esta estructura es coherente pues *el conflicto trabajo-familia*, que representa la necesidad de responder simultáneamente a las tareas laborales y a los trabajos domestico-familiares, está relacionado con las exigencias cuantitativas y el tiempo de las jornadas de trabajo (Moncada, Llorens et al., 2014; Moncada, Utzet et al., 2014). Esta misma estructura factorial fue observada en el AFE efectuado para la validación de la versión francesa del CoPsoQ (Dupret, Bocerean, Teherani, Feltrin y Pejtersen, 2012) y en la versión germana (Nuebling, Stöße, Hasselhorn, Michaelis y Hofmann, 2006).
- El tercer factor (F3) se ha denominado *Compensaciones del trabajo* y muestra una fiabilidad factorial de .80. Se compone de los ítems 21, 22, 23 y 24. Tiene que ver con el modelo teórico de esfuerzo-recompensa (Siegrist, 1996a, 1998), que indica que unas exigencias elevadas poco recompensadas, si son mantenidas en el tiempo, pueden convertirse en un estresor con posibilidad de dañar la salud del trabajador (HaSPA, 2012). La solución de este tercer factor coincide íntegramente con el factor teórico original y con la adaptación para España en su versión corta (Pejtersen, et al., 2010; Moncada, Llorens et al., 2014), midiendo *la inseguridad sobre las condiciones de trabajo* (ítems 21 y 22) y *la inseguridad en el empleo* (ítems 23 y 24). Estas subdimensiones recogen la incertidumbre laboral y el miedo a que las condiciones laborales empeoren.
- El cuarto factor (F4) se etiquetado como *Apoyo social y calidad de liderazgo*. Su fiabilidad factorial es de .94 y está compuesto por los ítems 17, 18, 19, 20, 25, 26, 27, 28, 29 y 30. A grandes rasgos, esta dimensión hace referencia a la ayuda, tanto material como emocional, que necesita recibir el trabajador de los demás miembros de la organización (HaSPA, 2012; Sauter et al., 1992). Coincide en parte con el modelo teórico de partida, midiendo *el conflicto de rol* (ítems 17 y 18), *la previsibilidad* (ítems 19 y 20) y *la calidad de liderazgo* (ítems 29 y 30). El factor original también integra *la claridad de rol*, sin embargo tal y como se ha mencionado en párrafos anteriores, en el presente AFE ha saturado en el factor 1. A este factor (F4) se suma una nueva dimensión, la de *capital social* -que está compuesta por las subdimensiones de *confianza vertical* (ítems 25 y 26) y *justicia* (ítems 27 y 28)-, y tiene que ver con la colaboración entre los diferentes miembros de la empresa para compartir recursos en su propio beneficio y en el de la organización. Aunque en el marco conceptual de origen es un factor independiente (Berthelsen et al., 2016), tiene relación con el apoyo social y la calidad de liderazgo, pues la base para esta colaboración se sostiene en la confianza entre los miembros, tanto vertical como horizontal (Moncada, Llorens, et al. 2014).

Para evaluar la simplicidad de esta solución factorial, se calculó el índice de simplicidad de Bentler (índice S, 1977), según el cual los valores cercanos a 1 muestran la máxima

simplicidad del modelo propuesto. El valor que se obtuvo fue $S=.99$, que indica una elevada simplicidad de la solución factorial obtenida. También se aplicó el índice de sencillez de carga *Loading Simplicity* (LS, Lorenzo-Seva, 2003): Cuando su valor es 1, indica que cada variable está relacionada con un único componente de la matriz de carga; es decir todos los valores de carga son cero excepto uno. Mientras que cuándo toma el valor de cero lo que muestra es que la comunalidad de cada variable está repartida uniformemente entre los factores; es decir, son valores idénticos y diferentes a cero (Lorenzo-Seva, 2003). El índice LS que se obtuvo es de .59 indicando que cada ítem está relacionado principalmente con un solo factor general. Como índice general de ajuste, se calculó la raíz media cuadrática residual (RMSR) que arrojó un valor de .05, que coincide con el umbral propuesto por el criterio de Kelly.

Para confirmar la estructura factorial de cuatro factores sugerida a partir del AFE en la primera submuestra, se estudió su replicabilidad en la segunda submuestra mediante un AFC (Brown, 2006; Ferrando y Anguiano-Carrasco, 2010). Al igual que se hizo para calcular el AFE, se aplicó el método de Ferrando et al. (2009) para eliminar la varianza debido al sesgo de DS. Seguidamente, con la matriz resultante se utilizó el criterio de mínimos cuadrados no ponderados *Unweighted Least Squares*, (ULS, Joreskog, 1977) recomendado por Ferrando y Anguiano-Carrasco (2010) debido a su eficiencia y porque no requiere de la estimación inicial de la varianza total (comunalidad).

Para poner a prueba el grado de ajuste al modelo de cuatro factores subyacentes, se ha utilizado una combinación de indicadores recomendados por la literatura (Bentler, 1990; Hu y Bentler, 1999; M.A. Ruiz, Pardo y San Martín, 2010; Tanaka, 1993), que no dependen del tamaño muestral, garantizando la idoneidad del modelo desde distintas perspectivas. Estos índices están basados en la evaluación de las correlaciones residuales, de modo que si se ha extraído el número de factores comunes apropiado, las correlaciones residuales tenderán a cero al quedar apenas varianza común (Lloret-Segura, et al., 2014):

- a) El índice de error cuadrático medio de aproximación *Root Mean Square Error of Approximation* (RMSEA, Steiger y Lind, 1980). Se trata de un índice inferencial de ajuste aproximado, utiliza el estadístico Chi-cuadrado (χ^2) y hace referencia a los grados de libertad. Está basado en los análisis de residuos y estima el error de aproximación del modelo factorial, revelando si es una propuesta razonable. Los valores que pueden considerar que el ajuste es excelente se sitúan por debajo de .05, mientras que los valores comprendidos entre .05 y .08 indicarían un ajuste aceptable, y los valores superiores a .08 estarían mostrando un ajuste inadecuado (Olobatuyi, 2006).
- b) El índice gamma o *Goodness of fit index-Adjusted* (GFI, Tanaka y Huba, 1989) es un indicador de bondad de ajuste absoluto que oscila entre 0 y 1, donde 1 es el ajuste perfecto. Este coeficiente indica la proporción de varianza total que puede ser explicada por el hipotético modelo que se está evaluando. Es equivalente al coeficiente de determinación múltiple en el modelo de regresión. Las cifras que estarían mostrando un buen ajuste del modelo, serían las superiores a .95 (Hu y Bentler, 1999; M.A. Ruiz et al., 2010; Tanaka, 1993).
- c) El último índice utilizado es el de ajuste comparativo *The Comparative Fit Index* (CFI, Bentler, 1990). Es un descriptivo que mediante la comparación de matrices de covarianzas, evalúa el grado de pérdida de ajuste al cambiar del modelo propuesto al

modelo nulo de variables incorreladas. Al igual que el anterior GFI, sus valores pueden variar entre 0 y 1, considerando que el ajuste sería óptimo si supera el .90, es decir, que al menos el 90% de la covarianza pueda ser reproducida por el modelo que se pone a prueba (Marsh, Hau y Grayson, 2005; Olobatuyi, 2006).

Como muestra la tabla 64, se puede concluir que el modelo de cuatro factores obtenido en el AFE tiene muy buen ajuste.

Tabla 64: Índices de bondad de ajuste del modelo sugerido por el AFE
Fuente: Elaboración propia

ÍNDICES DE BONDAD DE AJUSTE DEL MODELO		
RMSEA	GFI	CFI
.01	.98	.99

Por otra parte, también se aplicó el Coeficiente de congruencia de Tucker (1951) entre los patrones de carga en los factores obtenidos en el AFE y AFC, con el objeto de comprobar el grado de analogía o discordancia entre ambas estructuras. Los valores obtenidos fueron de .88, .90, .90 y .92 para cada factor, respectivamente, y la congruencia general fue de .90. Estos datos indican que la congruencia de las estructuras factoriales de las dos muestras son estadísticamente análogas, pues los valores del índice situados entre .85 y .94 apuntan que los factores pueden interpretarse como indicadores de similitud entre las dimensiones (Lorenzo-Seva y Ten Berge, 2006).

En consecuencia, dado que los estudios exploratorio y confirmatorio apuntan en la misma dirección al indicar que ambas muestras son representativas de la misma población, se decidió emplear el total de la muestra (N=563) para estimar la influencia de la DS en el CoPsoQ. El objetivo fue emplear el mayor número de participantes posible para obtener estimaciones más precisas, pues a mayor tamaño muestral mayor confianza se tendrá en la estabilidad del modelo propuesto (Lloret-Segura, 2014). Para ello, volvieron a utilizarse los análisis descritos en el método de Ferrando et al. (2009).

En la tabla 65 se muestran las cargas de los ítems en el factor de DS y en los cuatro factores de contenido después de la rotación oblicua para el conjunto de la muestra. Los resultados indican que la mayoría de ítems son indicadores relativamente puros, tal y como indicó el índice LS (Lorenzo-Seva, 2003). Esto es, cada ítem contiene una saturación alta en un único factor y baja en los demás factores. Como única excepción está el ítem 17 del CoPsoQ, que posee una saturación compleja al saturar por encima de .30 en más de un factor. En lo que respecta al sesgo de DS se puede comprobar como todos los marcadores de DS tienen saturaciones elevadas en el factor de DS, lo que permite calcular las cargas de DS en los factores de contenido. Siguiendo el criterio sugerido por Vigil-Colet et al. (2012), cargas factoriales superiores a .20 demuestran la posibilidad de presencia de sesgo debido a DS y cargas superiores a .30 indican un efecto sustancial de DS en el ítem.

Tal y como se recoge en la tabla 65, los ítems del CoPsoQ tienen una carga media en DS de .094 (es decir, excluyendo los ítems empleados como marcadores). Por lo tanto, en conjunto, no parecen afectados por la deseabilidad social. Sin embargo, un análisis detallado revela que

el 10% de los ítems del CoPsoQ tiene un posible efecto de DS. Concretamente los ítems con pesos factoriales iguales o superiores a .20 son los ítems 3, 5 y 18.

Tabla 65: Matriz de cargas factoriales del AFE controlando DS (n=563)

Fuente: Elaboración propia

Ítem	SD	Factor 1	Factor 2	Factor 3	Factor 4
DS1 +	0.61	0.00	0.00	0.00	0.00
DS2 +	0.62	0.00	0.00	0.00	0.00
DS3 +	0.55	0.00	0.00	0.00	0.00
DS4 +	0.36	0.00	0.00	0.00	0.00
CoPsoQ1 +	0.10	-0.09	0.67	-0.07	-0.04
CoPsoQ2 -	0.11	-0.02	0.64	-0.15	-0.11
CoPsoQ3 +	0.26	-0.04	0.57	-0.05	0.10
CoPsoQ4 +	0.19	-0.11	0.71	-0.01	0.04
CoPsoQ5 +	0.21	0.17	0.47	0.01	-0.02
CoPsoQ6 +	0.13	-0.02	0.65	0.20	0.04
CoPsoQ7 -	0.12	0.59	0.22	-0.03	0.16
CoPsoQ8 -	0.10	0.55	0.05	-0.02	0.04
CoPsoQ9 +	0.16	0.03	0.65	0.03	-0.07
CoPsoQ10 +	-0.02	0.19	0.74	0.04	0.06
CoPsoQ11 -	0.10	0.69	-0.01	-0.09	-0.11
CoPsoQ12 -	0.14	0.81	-0.06	-0.09	-0.18
CoPsoQ13 -	0.02	0.81	-0.04	-0.02	-0.01
CoPsoQ14 -	0.00	0.86	-0.02	0.02	-0.12
CoPsoQ15 -	-0.04	0.54	-0.08	0.09	0.18
CoPsoQ16 -	-0.04	0.52	-0.04	0.19	0.33
CoPsoQ17 +	0.17	-0.18	0.33	0.07	-0.35
CoPsoQ18 +	0.20	-0.12	0.27	0.09	-0.44
CoPsoQ19 -	-0.01	0.11	0.12	-0.15	0.56
CoPsoQ20 -	-0.10	0.15	-0.04	-0.11	0.61
CoPsoQ21 +	-0.03	0.05	0.15	0.53	-0.05
CoPsoQ22 +	-0.06	0.09	-0.01	0.72	-0.15
CoPsoQ23 +	-0.03	-0.04	-0.05	0.76	0.00
CoPsoQ24 +	0.04	-0.12	-0.03	0.61	-0.00
CoPsoQ25 -	-0.02	0.10	0.05	0.05	0.64
CoPsoQ26 -	-0.12	-0.04	-0.01	-0.08	0.82
CoPsoQ27 -	-0.08	-0.12	0.00	0.01	0.93
CoPsoQ28 -	-0.08	0.03	-0.09	0.01	0.78
CoPsoQ29 -	-0.05	0.02	-0.04	0.03	0.79
CoPsoQ30 -	-0.10	-0.02	-0.07	0.01	0.78
$r_{\theta\theta}=.86$.69	.88	.83	.80	.93

Se han sombreado en gris los ítems afectados por DS

\bar{X} = Media; σ = Desviación típica; As= Asimetría, K= curtosis

En un esfuerzo por demostrar si la presencia de DS puede generar diferencias entre grupos en los ítems del CoPsoQ, se optó por calcular las puntuaciones factoriales de los participantes en el factor de DS y dividirlos en dos grupos empleando la mediana como punto de corte. A

continuación, se compararon las medias de las saturaciones de DS de cada factor en ambos grupos (alta DS y baja DS), controlando y sin controlar el sesgo de discapacidad social. En la tabla 66 se recogen las medias para todos los factores expresadas en puntuaciones T (media 50 y desviación típica 10).

Tabla 66: Puntuaciones factoriales para los grupos bajos y altos en DS (n=563)
Fuente: Elaboración propia

	Grupo	F 1	F 2	F 3	F 4
Con DS	Baja DS	49.58	47.66	49.75	51.83
	Alta DS	50.42	52.34	50.25	48.18
Sin DS	Baja DS	49.33	49.16	49.92	50.64
	Alta DS	50.67	50.84	50.08	49.37

p < .001 (se destaca en sombreado gris la significación a 0.001)

p < .05 (se destaca en negrita la significación a 0.05)

El análisis multivariado de la varianza (MANOVA) sobre las puntuaciones factoriales sin eliminar el efecto de la DS muestra diferencias significativas entre los grupos ($F_{(4, 558)} = 13.606$; $p < .001$). En concreto, los análisis univariados fueron significativos para los factores de *Exigencias psicológicas en el trabajo* y *Apoyo social y calidad de liderazgo*. En el primer caso, el grupo alto en DS muestra mayores puntuaciones en la variable de *Exigencias psicológicas en el trabajo* que el grupo bajo en DS. Mientras que en la medida de *Apoyo social y calidad de liderazgo* el grupo alto en DS mostró puntuaciones menores que el grupo alto en DS. Por su parte, cuando los efectos de la DS se eliminan, el MANOVA pierde significación ($F_{(4, 558)} = 2.470$; $p < .05$). De hecho, la única diferencia a nivel univariado fue encontrada en el factor relativo a las *Exigencias psicológicas en el trabajo*. Los participantes con mayor DS mostraron mayores puntuaciones en *Exigencias psicológicas en el trabajo* que los participantes con menor DS.

En conjunto, los resultados indican que el CoPSoQ es parcialmente sensible a la DS y que, a pesar de estar afectados sólo un 10% de los ítems, la incidencia de la DS genera diferencias significativas entre grupos en aquellas dimensiones que engloban ítems con DS.

5.4.4. Funcionamiento Diferencial del ítem según el AN:

Para comprobar si el AN influye en las puntuaciones que miden condiciones laborales, se implementó el análisis del funcionamiento diferencial del ítem (DIF). Como primer paso se han calculado las puntuaciones del rasgo AN de los sujetos y se han dividido en dos grupos empleando la mediana como punto de corte, de modo que cada grupo tenga un número similar de informantes. Así, se obtuvo un grupo con puntuaciones altas en AN (grupo focal) y otro grupo con puntuaciones bajas en AN (grupo de referencia). El siguiente paso trata de analizar el DIF siguiendo el procedimiento basado en regresiones logísticas descrito por Hidalgo y Gómez (2006) e Hidalgo et al. (2005), para poder confirmar o no, la equidad de la escala.

Los resultados expuestos en la tabla 67 confirman la hipótesis de que los ítems del CoPSoQ no son equitativos, pues no funcionan de la misma forma entre los grupos AN alto y AN bajo. Como puede observarse el modelo 2 (M2 - M1) fue significativo para todos los ítems de la escala (significación menor o igual a .01), a excepción de los ítems 5, 7 y 22, por lo que se

evidencia la presencia de DIF uniforme. Por otra parte, ningún ítem del CoPsoQ mostró DIF no uniforme.

Tabla 67: Funcionamiento diferencial del ítem en AN (n=563)
Fuente: Elaboración propia

	M2-M1 Modelo 2	P	M3-M2 Modelo 3	p	R Nagelkerke (M3-M2)	R Nagelkerke (M2-M1)
CoPsoQ1 +	41.850	0.000	0.376	0.829	0.001	0.095*
CoPsoQ2 -	39.296	0.000	0.107	0.948	0.000	0.090*
CoPsoQ3 +	30.094	0.000	0.370	0.831	0.001	0.069
CoPsoQ4 +	52.037	0.000	0.261	0.878	0.001	0.117*
CoPsoQ5 +	8.070	0.018	0.037	0.982	0.000	0.019
CoPsoQ6 +	28.492	0.000	2.861	0.239	0.006	0.066
CoPsoQ7 -	7.087	0.029	0.141	0.932	0.001	0.016
CoPsoQ8 -	19.398	0.000	0.031	0.985	0.000	0.045
CoPsoQ9 +	54.000	0.000	1.120	0.571	0.002	0.122*
CoPsoQ10 +	20.996	0.000	0.271	0.873	0.000	0.049
CoPsoQ11 -	10.309	0.006	0.156	0.925	0.000	0.024
CoPsoQ12 -	15.908	0.000	0.046	0.977	0.000	0.037
CoPsoQ13 -	32.370	0.000	1.692	0.429	0.004	0.074*
CoPsoQ14 -	18.971	0.000	1.007	0.604	0.002	0.044
CoPsoQ15 -	27.949	0.000	0.251	0.882	0.001	0.064
CoPsoQ16 -	43.571	0.000	0.575	0.750	0.001	0.099*
CoPsoQ17 +	70.086	0.000	0.312	0.856	0.000	0.156*
CoPsoQ18 +	64.259	0.000	2.873	0.238	0.006	0.144*
CoPsoQ19 -	21.354	0.000	0.259	0.879	0.001	0.049
CoPsoQ20 -	52.119	0.000	1.758	0.415	0.003	0.118*
CoPsoQ21 +	16.633	0.000	1.022	0.600	0.002	0.039
CoPsoQ22 +	7.892	0.019	0.912	0.634	0.002	0.018
CoPsoQ23 +	12.051	0.002	2.993	0.224	0.007	0.028
CoPsoQ24 +	19.493	0.000	2.348	0.309	0.006	0.045
CoPsoQ25 -	44.784	0.000	2.824	0.244	0.006	0.102*
CoPsoQ26 -	57.566	0.000	1.411	0.494	0.003	0.129*
CoPsoQ27 -	61.971	0.000	4.193	0.123	0.009	0.139*
CoPsoQ28 -	63.281	0.000	0.028	0.986	0.001	0.141*
CoPsoQ29 -	50.810	0.000	0.267	0.875	0.000	0.115*
CoPsoQ30 -	37.911	0.000	0.112	0.946	0.000	0.087*

Se han sombreado en gris los valores con significación ≤ 0.01

* $\Delta R^2 > .070$

Como complemento a la detección del DIF se realizará una medición del tamaño de su efecto. La magnitud vendrá dada por las diferencias en el R^2 entre dos modelos. Se han seguido los criterios propuestos por Jodoin y Gierl (2001) que indican que si el ΔR^2 Nagelkerke es inferior a .035 el DIF sería inapreciable; si el valor se encuentra entre .035 y .070 sería moderado, y si es mayor a .070 el DIF sería elevado. Los resultados muestran que 15 ítems presentan DIF elevado, 10 ítems DIF moderado y 2 DIF inapreciable.

Una vez detectados los ítems con DIF se eliminaron sus puntuaciones del total del cuestionario. Tras la purificación se repitieron los análisis en todos los ítems con el total corregido (véase tabla 68).

Tabla 68: Funcionamiento diferencial del ítem purificado en AN (n=563)

Fuente: Elaboración propia

	M2-M1	p	M3-M2	p	R Nagelkerke (M3-M2)	R Nagelkerke (M2-M1)	Δ DIF
CoPsoQ1 +	39.292	0.000	0.159	0.924	0.000	0.089*	-0.006
CoPsoQ2 -	37.072	0.000	1.341	0.511	0.003	0.084*	-0.006
CoPsoQ3 +	25.214	0.000	0.586	0.746	0.002	0.057	-0.012
CoPsoQ4 +	45.066	0.000	2.767	0.251	0.006	0.101*	-0.016
CoPsoQ5 +	4.273	0.118	0.238	0.888	0.000	0.010	-0.009
CoPsoQ6 +	24.051	0.000	0.045	0.978	0.000	0.055	-0.011
CoPsoQ7 -	10.979	0.004	0.001	1.000	0.000	0.025	0.009
CoPsoQ8 -	20.521	0.000	0.260	0.878	0.000	0.047	0.002
CoPsoQ9 +	48.166	0.000	0.039	0.981	0.000	0.108*	-0.014
CoPsoQ10 +	16.171	0.000	2.798	0.247	0.006	0.037	-0.012
CoPsoQ11 -	15.457	0.000	0.267	0.875	0.001	0.035	0.011
CoPsoQ12 -	17.661	0.000	0.041	0.980	0.001	0.040	0.003
CoPsoQ13 -	26.526	0.000	0.152	0.927	0.001	0.060	-0.014
CoPsoQ14 -	16.860	0.000	0.153	0.926	0.000	0.039	-0.005
CoPsoQ15 -	21.834	0.000	0.452	0.798	0.001	0.050	-0.014
CoPsoQ16 -	32.654	0.000	0.129	0.938	0.000	0.074*	-0.025
CoPsoQ17 +	67.358	0.000	0.612	0.736	0.001	0.149*	-0.007
CoPsoQ18 +	61.651	0.000	0.416	0.812	0.000	0.137*	-0.007
CoPsoQ19 -	15.252	0.000	1.693	0.429	0.004	0.035	-0.014
CoPsoQ20 -	39.697	0.000	0.273	0.872	0.000	0.090*	-0.028
CoPsoQ21 +	12.339	0.002	0.025	0.988	0.000	0.028	-0.011
CoPsoQ22 +	3.463	0.177	0.823	0.663	0.002	0.008	-0.010
CoPsoQ23 +	7.129	0.028	3.248	0.197	0.008	0.016	-0.012
CoPsoQ24 +	15.713	0.000	0.147	0.929	0.000	0.036	-0.009
CoPsoQ25 -	32.481	0.000	0.294	0.863	0.000	0.074*	-0.028
CoPsoQ26 -	43.414	0.000	1.231	0.540	0.002	0.098*	-0.031
CoPsoQ27 -	45.873	0.000	1.313	0.519	0.003	0.103*	-0.036
CoPsoQ28 -	49.345	0.000	0.003	0.999	0.000	0.110*	-0.031
CoPsoQ29 -	37.682	0.000	0.438	0.803	0.001	0.085*	-0.030
CoPsoQ30 -	28.715	0.000	0.707	0.702	0.002	0.065	-0.022

Se han sombreado en gris los valores con significación ≤ 0.01 y en la última columna los valores con Δ del DIF uniforme

* $\Delta R^2 > .070$

Al realizar la purificación del total, la mayoría de ítems redujeron el DIF detectado, salvo 4 ítems en los que aumentó el DIF uniforme. Sin embargo, los mismos ítems en los que se había identificado DIF significativo al 0.01, nuevamente volvieron a mostrarlo, a excepción del ítem 23. El ítem CoPsoQ7 una vez purificado muestra DIF significativo, aunque su nivel es inapreciable. Con respecto a la magnitud del efecto las variaciones son poco importantes. Ahora los resultados observados son 2 ítems con DIF inapreciable, 12 ítems con DIF moderado y 13

ítems con DIF elevado, es decir tan sólo se reducen dos ítems con DIF elevado, pasando a DIF moderado.

Estos resultados confirman la presencia de DIF considerable en los ítems del CoPsoQ. Para saber si el DIF se muestra en todas las categorías de respuesta o sólo en algunas, se aplicó el análisis de categorías acumuladas de Mellenbergh (1995), cuyos resultados corroboran el DIF en todas las opciones de respuesta (de 0 a 4), es decir, sujetos bajos y altos en AN, igualados en el nivel de exposición a los factores psicosociales, responden con diferente frecuencia a las distintas alternativas de respuesta de la mayoría de los ítems del CoPsoQ.

Por último, con objeto de averiguar la dimensión de las diferencias originadas por el DIF uniforme en los ítems de la escala CoPsoQ, se llevó a cabo un MANOVA donde el AN se incluyó como factor fijo y como variables dependientes los factores limpios de sesgo DS del CoPsoQ. Los resultados indican nuevamente la existencia de diferencias significativas entre los grupos de alto AN y bajo AN, incluso tras haber corregido el sesgo de DS ($F_{(4, 558)} = 26.209$; $p < .001$). En la tabla 69 se presentan las medias para todos los factores expresadas en puntuaciones T (Media = 50; DT = 10). Como era de esperar, todos los análisis univariados fueron significativos, mostrando los sujetos con alto AN mayores puntuaciones en *Exigencias psicológicas en el trabajo* (F2) y en *Compensaciones del trabajo* (F3), y menores puntuaciones en *Control sobre el trabajo* (F1) y *Apoyo social y calidad de liderazgo* (F4), que los sujetos bajos en AN.

Tabla 69: Puntuaciones factoriales para los grupos bajos y altos en AN (n=563)
Fuente: Elaboración propia

Grupo	F 1	F 2	F 3	F 4
Bajo AN	51.62	47.28	48.58	52.99
Alto AN	48.40	52.69	51.40	47.04

$p < .001$ (se destaca en sombreado gris la significación a 0.001)
 $p < .05$

5.5. DISCUSIÓN DE LOS RESULTADOS Y CONCLUSIONES

5.5.1. Discusión de resultados y conclusiones sobre el control de la DS

En el presente estudio se han empleado diferentes técnicas para disminuir el potencial impacto de la DS, tanto antes como después de la recolección de datos. En la fase previa a la recogida de información se aplicaron las siguientes pautas: i) garantizar la confidencialidad y el anonimato (Lajunen y Summala, 2003; Morales, 2006; Paulhus y Reid, 1991; Podsakoff et al., 2003), ii) demandar específicamente la sinceridad de respuestas (Gordon, 1987; Podsakoff et al., 2003), y iii) administrar los test vía internet (Risko et al., 2006). Estos procedimientos pudieron haber contribuido a minimizar en gran medida las respuestas socialmente deseables.

Una vez obtenida la información se aplicó la técnica de análisis factorial de Ferrando, (2005) y Ferrando et al. (2009) para conseguir un factor específico de DS. Al emplear esta técnica se ha podido corroborar, al menos en parte, la hipótesis H1 “*La medida de los factores psicosociales de riesgos estará sesgada debido al efecto de la deseabilidad social*”, pues el 10% de los ítems afectados por DS inflan las puntuaciones de exposiciones adversas a riesgos psicosociales generando diferencias estadísticamente significativas. En concreto, los análisis

efectuados señalan que hay tres reactivos afectados por DS -en concreto, los ítems 3, 5 y 18-, y otro ítem muy cerca del límite de .20 (el ítem 4).

Los ítems 3 “¿Con que frecuencia hay momentos en los que necesitarías estar en la empresa y en casa a la vez?”, 4 “¿Con qué frecuencia sientes que tu trabajo te ocupa tanto tiempo que perjudica a tus tareas doméstico-familiares?” y 5 “¿Con que frecuencia en el trabajo tienes que ocuparte de los problemas personales de otras personas?” pertenecen al factor etiquetado como *Exigencias psicológicas*. Concretamente los dos primeros integran la subdimensión de doble presencia según el modelo teórico original. Por su parte, el ítem 18 “¿En qué medida tienes que hacer tareas que crees que deberían hacerse de otra manera?” forma parte del factor conceptualizado como *Apoyo social y calidad de liderazgo*. Precisamente, es uno de los ítems que evalúa el conflicto de rol.

La deseabilidad social encontrada en los dos primeros ítems apunta en la misma dirección que otras investigaciones (ej., Achen y Stafford, 2005, Kan, 2008; Press y Townsley, 1998; Sigelman et al., 1981), donde los hombres, en comparación con las mujeres, mostraron discrepancias al alza cuando informaban sobre su participación en las tareas domésticas. De hecho, es un estudio previo con la versión corta del CoPsoQ-Istas21 v.1.5 (Louzán, 2014), el factor de doble presencia mostró los niveles de exposición más desfavorables para la salud (45%), a pesar de ser obtenido en una muestra integrada fundamentalmente por hombres.

De acuerdo con la redacción del ítem 5, sería socialmente deseable responder de modo afirmativo, pues los respondientes estarían dando valor a su trabajo al señalar que sus funciones van más allá de las tareas propias de su puesto de trabajo. En esta misma línea apuntan, Gamero-Burón y González (2013) que, en su investigación sobre las altas puntuaciones de estrés recogidas en diferentes encuestas, como la Encuesta nacional de salud (ENSE, 2006) y el *International Social Survey Programme* (ISSP, 1997), indicaron la posibilidad de que los resultados estuviesen distorsionados por DS. Una posible explicación reside en que los trabajadores podrían estar motivados a mostrar que experimentan estrés para evidenciar que son miembros valiosos de la organización.

Respecto al ítem 18, cabe señalar que los trabajadores tienen sus propias ideas sobre el papel o rol que deben desempeñar en su puesto de trabajo (de Arquer, Daza y Nogareda, 1995, NTP 388). En este sentido, lo socialmente deseable sería indicar que ellos son los que saben cómo tienen que hacerse las tareas, pues estaría relacionado con un buen conocimiento del puesto de trabajo y con tener las habilidades para desarrollarlo.

Al examinar la escala en conjunto, los análisis realizados para comprobar si los trabajadores con puntuaciones altas en DS mostraban diferencias al compararlos con los sujetos con DS baja, se ha observado que cuando no se elimina el sesgo de DS existen diferencias significativas en los factores de *exigencias psicológicas* y *apoyo social y calidad de liderazgo*. Concretamente el grupo alto en DS reporta mayores puntuaciones en *exigencias psicológicas* y menores en *apoyo social y calidad de liderazgo* en comparación con los sujetos con DS baja. Aquí, es necesario señalar que la variable de *apoyo social y calidad de liderazgo* funciona a la inversa, es decir, son las puntuaciones menores las que indican mayor exposición al riesgo, lo contrario que con la variable de *exigencias psicológicas*. También es preciso tener en cuenta que son justamente las dimensiones que contienen los ítems afectados por DS. De hecho, cuando se elimina el efecto del sesgo de DS, se reduce notablemente la significación estadística,

encontrando sólo diferencias en las puntuaciones referentes a las exigencias psicológicas. Este es un hallazgo importante, pues si se tienen en cuenta los resultados de los cuestionarios sin controlar la DS, su interpretación llevaría aparejado implementar medidas preventivas correctoras de mayor magnitud que si se toman en cuenta los datos resultantes una vez depurado el sesgo.

Aunque en los análisis realizados se ha asumido la DS como varianza error y no como rasgo de personalidad -pues no existe un criterio unánime al respecto-, lo cierto es que en cualquiera de estas posturas la DS es un obstáculo que dificulta la obtención de medidas puras de la actitud o rasgo que se pretende medir. Esto hace necesario utilizar técnicas encaminadas a disminuir sus efectos o por lo menos a poder determinar la importancia del sesgo en la escala. (Ver apartado 4.3.2 Formas de control de la deseabilidad). Lo ideal es controlar el sesgo en la fase del diseño de las escalas, por ejemplo, construyendo ítems socialmente neutrales (De las Cuevas y González de Rivera, 1992; Hough et al., 1990; Morales, 2006). No obstante, cuando la escala ya está creada, pueden aplicarse otras medidas de control con objeto de minimizar la DS.

En consecuencia, aunque se han tomado medidas para disminuir la DS (ej., datos extraídos online), los análisis efectuados indican que el cuestionario CoPsoQ-ISTAS21 v.2 es parcialmente sensible a la DS y que, a pesar de estar afectados sólo un 10% de los ítems, la incidencia de la DS genera diferencias significativas entre grupos. Estos datos apoyarían los argumentos de Duro (2005), quien afirma que (1) las medidas autoinformadas de variables de ámbito psicosocial laboral son sensibles a DS, y (2) que la distorsión sería positiva o negativa dependiendo de si la naturaleza de la variable favorece o desfavorece al trabajador.

Los factores más sensibles a la DS han sido las *exigencias psicológicas* (factor 2) y *el apoyo social y calidad de liderazgo* (factor 4), en comparación con el resto de factores. Como ejemplo, puntuar alto en el ítem del factor 2 “ *tienes que trabajar muy rápido*” y puntuar bajo en el ítem del factor 4 “*se puede afirmar que tu jefe inmediato planifica bien las tareas*”, formaría parte de los estereotipos sociales desde el punto de vista del trabajador. Así, trabajar con rapidez o bajo presión sería la respuesta más deseable debido a sus connotaciones positivas, pues implicaría mayor productividad al poder realizar más tareas en menor tiempo, cualidad que se sabe es muy valorada por los empleadores (Sánchez-Díaz, 2013). Por otra parte, que un trabajador sea capaz de llevar a cabo sus tareas, a pesar de existir una mala planificación de las mismas, está relacionado con una mayor eficiencia y puede asociarse con las competencias de manejo y resolución de problemas (Alles, 2008).

En vista de los datos obtenidos, puede afirmarse que el sesgo DS influye en los resultados de la escala actuando como distorsionador de las puntuaciones y que, por lo tanto, las interpretaciones e inferencias realizadas sin controlar la tendencia de respuesta tendrían que ser tomadas con cautela.

5.5.2. Discusión resultados y conclusiones sobre la detección de DIF

Los análisis efectuados según el método propuesto por T.R Miller y Spary (1993) revelan que no es posible realizar una comparación de puntuaciones directas entre sujetos con AN alto y AN bajo, pues la gran mayoría de los ítems presentan DIF en contra del grupo AN alto. Esto es, igualados ambos grupos al mismo nivel de exposición de factores de riesgo psicosocial, la

probabilidad de respuesta del grupo AN alto presenta mayores puntuaciones en el sentido de la escala y menores en el sentido inverso, en todos los niveles de la variable.

Dicha información es corroborada con el análisis multivariado de la varianza (MANOVA), donde, aún corregido el sesgo de DS, se presentan diferencias significativas entre el grupo de alto AN y bajo AN. Concretamente, el grupo con alto AN presenta peores exposiciones en los cuatro factores de riesgo, mostrando puntuaciones más altas en los factores en la misma dirección de la escala (*exigencias psicológicas y compensaciones del trabajo*) y mostrando puntuaciones inferiores en los factores con funcionamiento inverso (*control sobre el contenido del trabajo y apoyo social y calidad de liderazgo*).

Según esto, los sujetos con peor estado de ánimo (alto AN) estarían percibiendo exposiciones más desfavorables en comparación con los sujetos más calmados y serenos (baja AN), lo que confirma la hipótesis H2: *La medida de los factores psicosociales de riesgos contiene ítems con DIF en contra de los trabajadores que presentan alto afecto negativo*.

Los índices estadísticos de DIF muestran un hipotético sesgo, que puede verse confirmado si se determina la existencia de interacción entre el contenido de los ítems y alguna característica del grupo que contamine el proceso de evaluación. La ratificación puede encontrarse en los postulados derivados de la literatura científica que sostienen que un trabajador que presenta alto AN tiene una visión pesimista del mundo que lo rodea (J. Martín et al., 2007), y con mayor frecuencia se siente menos satisfecho y percibe estrés de mayor intensidad (Burke, 2001; Magnus et al., 1993; Sandi y Calés, 2000). De esa forma estaría percibiendo los factores de riesgo psicosocial de modo más perjudicial que los sujetos que no muestren ese rasgo.

La información obtenida indica que los sujetos con alto AN presentan las peores exposiciones, confirmando así los resultados de otros estudios, donde la explicación de las diferencias de percepción entre ambos grupos (alto AN y bajo AN) vendría marcada por la personalidad, que actúa como variable confusora al intervenir en la percepción de los estresores laborales (Caicoya, 2004; J. Martín et al., 2007; Staw, et al., 1986). Concretamente, los hallazgos derivados de este estudio corroboran las evidencias existentes acerca de que los sujetos con puntuaciones altas en AN reportan mayores amenazas de estrés (Forgas, Chan y Laham, 2001; Johns y Saks, 2005; J. Martín et al., 2007; Sandi y Calés, 2000; Suls, 2001) y de mayor intensidad (Burke, 2001; Magnus et al., 1993; Sandi y Calés, 2000) que los sujetos con un mejor estado de ánimo (bajo AN) (Chay, 1993), al percibir peores condiciones de trabajo y menor apoyo social (Brief et al., 1988; Elliot et al., 1994; Forgas y George, 2001; Levin y Stokes, 1989).

Los datos encontrados en esta investigación son contundentes. Indican DIF significativo y uniforme en el 90% de los ítems de la escala CoPsoQ; es decir, se han obtenido puntuaciones mayores o menores de las que realmente deberían ser para el grupo categorizado con alta afectividad negativa, poniendo en entredicho la validez de constructo de las mediciones. Estos hallazgos contribuirían a dotar de fundamento las sospechas de la EU-OSHA (2005) sobre el papel distorsionador del AN en la percepción que tienen los trabajadores sobre sus condiciones laborales. Sumándose así, a muchos otros que responsabilizan al AN como potencial fuente de distorsión de los resultados (Caicoya, 2004, F. Jones y Kinman, 2001; Razavi, 2001). Pero sobre todo, estarían confirmando las conclusiones de los estudios que aseguran que las medidas de

autoinforme que evalúan estrés o estresores son susceptibles de contener varianza explicada por AN, que inflaría los resultados (Brief et al., 1988; Watson y Pennebaker, 1989).

En ese orden de ideas, se hace necesario tomar medidas para disminuir la varianza explicada por el AN en las evaluaciones de riesgo psicosocial. Así, EU-OSHA (2005) sugiere principalmente la triangulación de datos, que consiste en contrastar datos cuantitativos con cualitativos (Caicoya, 2004; Cox et al., 2000; F. Jones y Kinman, 2001). Sin embargo, lo más recomendado en la literatura, especialmente cuando lo que se evalúan son estresores o salud de manera subjetiva, es incluir medidas específicas de AN en los test. De esta manera se puede controlar estadísticamente su efecto (Razavi, 2001), y además, daría la posibilidad de evaluar a los trabajadores separadamente según grupos de similitud (ej., AN alto) para posteriormente comparar si sus puntuaciones son similares y poder disminuir el factor de confusión en las medidas.

5.5.3. Implicaciones teóricas y prácticas

La principal contribución de este segundo estudio es el cuestionamiento de la validez de las medidas de autoinforme que evalúan factores psicosociales de riesgo en el trabajo. Los resultados despejan el vacío de conocimiento existente al respecto, al corroborar que los ítems están altamente influenciados por el rasgo de AN y, en menor medida, por la DS. La existencia de sesgos en estas evaluaciones implica tener que tomar con cautela los resultados de aquellas evaluaciones que no hayan tomado medidas para su control. Por otro lado, la confirmación de que los trabajadores caracterizados por alto AN experimentan en mayor grado los estresores puede tomarse como referencia para la reformulación de las fundamentaciones teóricas en la que se sustentan la mayoría de los test de evaluación psicosocial de origen laboral.

Por su parte, las implicaciones prácticas derivadas de este estudio apuntan a la necesidad de concienciar a la comunidad científica y profesional sobre la importancia de los sesgos de DS y AN en los ítems de autoinforme que miden condiciones laborales y sobre las consecuencias de infravalorarlos. En ese sentido, los hallazgos derivados de esta tesis pueden contribuir a mejorar las interpretaciones derivadas de los autoinformes de evaluación de riesgos psicosociales. A su vez, estos resultados pretenden ser de ayuda en el avance metodológico de las evaluaciones de riesgos psicosociales al promover el (re)diseño de nuevas herramientas de medición que controlen desde su gestación el impacto de la DS y el AN, o también en la selección de ítems adecuados dentro de las herramientas preexistentes. Los resultados obtenidos, también pueden sentar las bases para la confección de un baremo que tenga en cuenta la distinta percepción de los trabajadores con alta DS y alta AN.

A la espera de que esto suceda, los profesionales e investigadores interesados pueden controlar los sesgos de respuesta y el afecto negativo de los cuestionarios de evaluación psicosocial, (a) siguiendo alguna de las propuestas presentadas en el marco teórico de esta tesis, o (b) replicando la metodología utilizada en la parte empírica. Esta última aplicación práctica lleva aparejada la necesidad de contar con conocimientos avanzados de psicometría, un campo poco utilizado en la mayoría de evaluaciones de riesgos psicosocial que se llevan a cabo en nuestro país. En este aspecto, las organizaciones y organismos preocupados por la prevención de riesgos deberían interesarse en proporcionar las herramientas necesarias para llevar a cabo

esta labor, entre ellas se encontrarían programas formativos especializados en análisis de datos multivariantes.

5.5.4.Limitaciones y líneas de investigación futuras

Esta investigación ha arrojado información de gran importancia y utilidad para el campo de la prevención de riesgos psicosociales de origen laboral, sin embargo, no está exenta de limitaciones, las cuales sería conveniente salvar en futuros estudios que sigan las líneas de investigación aquí expuestas.

La primera limitación reside en la propia naturaleza subjetiva de los datos recabados. Es una limitación común a toda investigación realizada exclusivamente con autoinformes, por lo que sería recomendable utilizar medidas multimétodo. No obstante, teóricamente, es muy complicado evaluar las condiciones laborales adversas sin contar con la propia percepción del trabajador, sobre todo si se tiene en cuenta que las evaluaciones de riesgos psicosociales evalúan hechos probables y no probados. La varianza de método común es otra de las debilidades asociadas a los cuestionarios de autoinforme, por estar éstos influidos por factores subjetivos, estados de ánimo y diferencias individuales, entre otras circunstancias (Littlewood, 2006). Para superar este sesgo, se han llevado a cabo diversas técnicas con objeto de detectarlo y minimizarlo (véanse epígrafes 5.3.6.1 y 5.4.1), como por ejemplo el cálculo del test de Harman, cuyo resultado ha mostrado que este sesgo no limita la actual investigación.

Siguiendo con las limitaciones que afectan a los estudios basados en medidas de autoinforme, se encuentra la imposibilidad de los encuestados de aportar información adicional por ser cuestionarios de respuesta cerrada. Esta cuestión impide generalizar a variables no observadas en el estudio. Aunque en el cuestionario presentado a los participantes se ha incluido un apartado final de respuesta abierta para observaciones y consideraciones, sería recomendable que futuros trabajos pudiesen incluir la triangulación de medidas para mejorar la robustez metodológica.

Respecto a la muestra utilizada, al proceder de un muestreo no probabilístico, hace que se cuestione su representatividad, lo que impide generalizar con seguridad los resultados a la población (R. Hernández et al., 2014). Intentando conseguir la máxima representatividad posible se optó por realizar un muestreo no aleatorio por cuotas balanceado por sexo, debido a su semejanza al muestreo probabilístico estratificado. Las cuotas utilizadas de referencia pertenecen a la última publicación EPA (INE, 2016). A pesar de todo, es difícil saber si los resultados obtenidos podrían estar sesgados por las características de la muestra. Por otro lado, al ser una muestra incidental en la que se reclutan participantes hasta completar determinadas cuotas, ésta es susceptible al sesgo de selección y, en consecuencia, cabe plantearse si las motivaciones de los que optaron por participar son distintas a los que rechazaron formar parte del estudio. Por ejemplo, está presente el interrogante de si existe interés por parte de los que aceptaron participar en sobreestimar o subestimar los riesgos a los que están expuestos en sus respectivos puestos de trabajo. Futuros trabajos que pretendan replicar o mejorar este estudio deberían utilizar un muestreo aleatorio para disipar posibles sobre la representatividad de la muestra.

Por otra parte, la toma de datos se realizó vía web que, paradójicamente, aunque permite ampliar el tamaño muestral y minimizar el sesgo de deseabilidad social (Risko et al., 2006),

contribuye a restar representatividad a los resultados obtenidos debido a que sólo pueden rellenar el cuestionario las personas con acceso a internet. Otra de las limitaciones referentes a la muestra y derivadas del sistema de reclutamiento, tiene que ver con la pertenencia de los encuestados a distintas y variadas organizaciones. Al no pertenecer los trabajadores a la misma empresa, no necesariamente están expuestos a las mismas condiciones laborales, y por ende no se puede garantizar que, realmente, no existan otras variables que puedan explicar los datos encontrados, por lo que tampoco se puede comparar la percepción de las condiciones laborales entre sujetos de la misma organización. Para comprobar si el procedimiento de recogida de información ha influido en las respuestas, se recomendaría a trabajos venideros que quieran continuar con la línea de investigación que repliquen la metodología con muestras presenciales de lápiz y papel; de ese modo también podría valorarse si es mayor el nivel de sesgo encontrado de DS. A su vez, sería conveniente utilizar como muestra trabajadores de la misma organización y, siempre que sea posible, utilizar la triangulación metodológica.

Debido a las restricciones de tiempo inherentes a los estudios de doctorado regulados por la reciente normativa (RD 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado), ha tenido que optarse por un estudio de diseño transversal, lo que supone una medida de las variables en un solo momento de tiempo, es decir, una *foto fija* que imposibilita deducir cualquier tipo de causalidad. Tanto las actitudes como las condiciones laborales son susceptibles de variar con el paso del tiempo y por lo tanto alterar los datos aquí obtenidos. Para superar esta salvedad, se sugiere como reto futurible realizar estudios de corte longitudinal para aumentar la fiabilidad de los resultados y poder establecer relaciones de causalidad. Concretamente, se contribuiría a identificar con mayor claridad cuál es el papel que cumple la AN en la relación trabajo-estrés (Oliver, Mansell y Jose, 2010).

En lo concerniente a la medición de la variable del afecto, una de las limitaciones encontradas alcanza al uso del instrumento de medida PANAS. Para esta investigación se han evaluado los afectos de manera general y en un determinado momento. Sería recomendable que en futuros trabajos se utilizase la herramienta para evaluar los afectos de manera prolongada en el tiempo, de ese modo podría determinarse con mayor exactitud si el AN es consecuencia de una emoción provocado por un estímulo concreto, como es la exposición a factores de riesgo adversos o si, por el contrario, es un rasgo estable no relacionado con el estímulo (Padrós, Soria-Mas y Navarro, 2012).

Asimismo, en lo concerniente al DIF sobre AN alto y AN bajo, es preciso aclarar que su estudio sobrepasa el ámbito de la prevención de riesgos laborales al amparo de la LPRL, pues se está evaluando al individuo y no las condiciones laborales. Por consiguiente, y aunque se recomienda la inclusión de medidas de AN al evaluar estresores (Razavi, 2001) para evaluar de una manera más controlada su efecto en la percepción de las condiciones laborales, es preciso tener en cuenta que esta medida solo debe de utilizarse de manera complementaria (ITSS, 2012), pues va más allá de las competencias relativas a la prevención de riesgos laborales. Atendiendo a esta realidad, lo recomendable sería profundizar sobre el diseño de nuevas escalas de evaluación de condiciones laborales que tengan en cuenta el posible DIF respecto al AN, en aras de perfeccionar la selección de los ítems.

Como sugerencia, resultaría fructífero para la práctica profesional explorar en la misma línea otras herramientas de evaluación psicosocial (ej., FPSICO). Así, con los resultados de

unas y de otras, podrían realizarse baremos de percepción psicosocial adversa para grupos con alto DS y alta AN.

En lo que concierne a los técnicos de prevención de riesgos laborales (no psicólogos), según su perfil formativo, lo habitual es que no estén familiarizados con la teoría de los test, ni con conocimientos relativos a técnicas de construcción y uso de los test, ni con métodos estadísticos. Esto hace que la administración de los cuestionarios de evaluación se limite a aquellos que puedan aplicarse, corregirse e interpretarse con ayuda de un manual. Ello dificulta en gran medida la posibilidad de detección y corrección de sesgos de respuesta (ej., DS) utilizando técnicas estadísticas. En concreto, los análisis efectuados en este estudio para conseguir la medida pura de DS requieren de conocimientos avanzados en estadística. Para salvar este escollo, se recomienda la utilización de un software sencillo como, por ejemplo, el “*Psychological Test Toolbox*” (Navarro, Vigil-Colet, Ferrando y Lorenzo-Seva, 2016), el cual facilita la obtención de (1) la matriz de correlaciones libres de sesgo (aquiencia y deseabilidad social) para utilizar tanto en análisis factorial exploratorio como en el confirmatorio; (2) las estimaciones factoriales exploratorias al haber sido eliminados los sesgos de respuesta; y (3) las puntuaciones factoriales libres de sesgos.

Finalmente, cabe señalar que aunque la actividad investigadora sobre evaluación de riesgos psicosociales es amplia, también está desequilibrada en cuanto a temáticas a explorar. La falta de producción científica relativa a las amenazas inherentes a los autoinformes de evaluación psicosocial, dificulta la posibilidad de comparación de resultados entre estudios. Tal vez tenga que ver el proceso de formación de los técnicos de prevención, y en ese sentido sería necesario ahondar en la ciencia psicométrica. Pues técnicas de análisis de datos más sofisticadas, como las regresiones logísticas o las ecuaciones estructurales, contribuyen a incrementar la significación de los resultados y, por consiguiente, permitirían un mayor avance de conocimiento sobre este campo. Estos motivos hacen necesario por parte de las instituciones responsables, fomentar y desarrollar programas de investigación más específicos sobre el análisis científico en las evaluaciones de riesgo psicosocial.

A pesar de las limitaciones mencionadas, esta investigación sirve de apoyo para corroborar el papel de los sesgos de respuesta en la validez de los resultados, así como el de la AN en la percepción que tiene el trabajador en las condiciones laborales. Debido a la falta de estudios en este ámbito, los hallazgos de aquí derivados podrían servir de punto de partida para futuras investigaciones que decidan profundizar en la relación de ambas variables con las medidas de autoinforme. Serían bienvenidos estudios que contribuyan a ampliar esta línea de investigación, tanto refutando como mejorando las soluciones metodológicas aquí aportadas.



6. CONCLUSIONES

6.1. CONCLUSIONES FINALES

En estos últimos años los factores de riesgo psicosocial están siendo objeto de profunda reflexión y debate, tanto por la preocupante influencia que ejercen sobre la salud y el bienestar de los trabajadores como por las dificultades inherentes a su prevención. Fruto de la discusión generada nació esta tesis, y entre sus objetivos estaba el deseo de dar visibilidad a las amenazas que actúan de lastre a la hora de evaluar y prevenir los riesgos psicosociales. Para ello, se realizó una revisión de la literatura disponible relacionada con las amenazas a la validez de las mediciones efectuadas en la evaluación de los riesgos psicosociales, de la cual surgieron interrogantes que era necesario investigar y contrastar utilizando el rigor científico que, a su vez, contribuyeron a justificar la necesidad de esta tesis (véase epígrafe 4.3.1.).

Para llevar a cabo este proyecto se diseñó una investigación desde un enfoque mixto, cualitativo - cuantitativo. Los trabajos de investigación cualitativa se realizaron en primer lugar, y se centraron en sondear la opinión de un grupo de distintos expertos acerca de las dificultades inherentes a las evaluaciones de riesgos psicosociales con cuestionarios de autoinforme. Para cuantificar empíricamente el impacto de algunas de las principales amenazas descritas por dichos expertos (DS) y comprobar la influencia del AN, se diseñó la investigación cuantitativa.

El estudio cualitativo ha sido fruto del juicio y consenso entre estos expertos, a partir de los cuales se ha conseguido establecer un escenario real sobre las dificultades de evaluar riesgos psicosociales, siendo posible poner nombre a las debilidades de las que adolecen este tipo de evaluaciones, a la vez que se recopilaba una batería de recomendaciones para fortalecer y mejorar su validez. Este procedimiento ha arrojado información relevante, especialmente, sobre los sesgos de respuesta de las medidas de autoinforme destinadas a evaluar las condiciones de trabajo, por ser éste el foco de mayor controversia, y por la falta de evidencias científicas al respecto. Con ese fin se diseñó el segundo de los estudios, ésta vez cuantitativo. Los objetivos a alcanzar han sido, por un lado, determinar si la DS tiene algún efecto en estas medidas y, por otro lado, analizar el funcionamiento diferencial del ítem para analizar posibles sesgos producidos por el rasgo de AN. A continuación se muestran algunos de las conclusiones más notables obtenidas de los resultados en conjunto de ambos estudios:

1. Subestimación de los riesgos psicosociales:

En primer lugar, los expertos se pusieron de acuerdo para afirmar que los riesgos psicosociales son subestimados por las empresas, confirmando que se les presta menor atención y recursos para su prevención y gestión. Es difícil rebatir este enunciado en vista de la cantidad de datos estadísticos que lo avalan, como los obtenidos de las encuestas, tantas veces mencionadas, ESENER (2010) y ESENER-2 (2015). Circunscribiéndonos a la realidad española, estas encuestas muestran que solamente el 15% de empresas que cuentan con más de 10 trabajadores toman medidas para reducir el estrés. La justificación de algunos directivos (42%) se basa en que es más difícil gestionar los riesgos psicosociales que otro tipo de riesgos

(ESENER, 2010). El poco peso que poseen los riesgos psicosociales en prevención se deduce notoriamente de la baja demanda de profesionales en el sector. Así, según ESENER-2 (2015), en España un 76% de los servicios de prevención, tanto ajenos como propios, solicitan los servicios de expertos en prevención de accidentes laborales, frente a un pobre 15.70% de profesionales del área psicosocial. Esto es así porque, lamentablemente, los riesgos psicosociales apenas son incluidos en las evaluaciones iniciales de riesgos y la mayor parte de las veces que se analizan estos riesgos es por requerimiento de la inspección laboral para evitar multas y sanciones (C. Molina, 2010).

A la hora de buscar culpables, son muchas las circunstancias mencionadas, como por ejemplo, las indicadas por Meliá (2008): la novedad de los riesgos psicosociales, la falta de formación especializada en psicometría y psicología del trabajo de muchos profesionales, o la calidad de los instrumentos de medida. Peiró (1999a) es otro que se suma en apuntar directamente a la disponibilidad y calidad de las herramientas de medición. Las causas que dificultan las evaluaciones de riesgo psicosocial pueden ser muchas y de distinta naturaleza, pero sin duda, se lleva la palma la falta de cultura preventiva. Nada menos que el 90.50% de empresas españolas encuestadas aseguraron que los principales motivos que les movían a gestionar los riesgos psicosociales son cumplir la legislación y evitar las multas y sanciones impuestas por la inspección de trabajo (ESENER-2, 2015).

Esta ausencia de cultura en prevención de riesgos psicosociales lleva aparejada un reiterado incumplimiento de la LPRL, poniendo en peligro la salud y bienestar de los trabajadores. Por tanto, sería beneficiosa una actuación inspectora de mayor calado, así como plantearse una renovación de la mencionada LPRL, que, tras llevar más de 20 años en vigor y atendiendo a las nuevas realidades organizacionales y económicas (Eurostat, 2017), sobrevuela sobre ella la obsolescencia, especialmente en lo referente a los riesgos psicosociales. Precisamente, ni siquiera son objeto de mención en dicha ley de manera explícita, ni se dispone de normas UNE o de guías técnicas específicas que orienten sobre su evaluación, dejando al libre albedrío de los responsables en prevención la elección de las herramientas de medición.

2. Falta de consenso sobre la mejor herramienta de evaluación a utilizar:

Tradicionalmente, las respuestas al estrés han sido valoradas utilizando cuatro grupos de herramientas diferenciadas (a saber, medidas de autoinforme, conductuales, fisiológicas, bioquímicas), siendo las medidas de autoinforme las más utilizadas (Lester, Nebel y Baum, 1994) y las únicas disponibles dentro de las competencias de los técnicos de prevención de riesgos laborales. En este sentido, un 87.50% de los expertos participantes en el primer estudio recalcó la importancia de la autopercepción del trabajador para valorar los riesgos psicosociales. Sin el punto de vista del trabajador, sería difícil determinar el grado de exposición a los factores de riesgo psicosocial (Moreno y Báez, 2010; Peiró, 2008; Villalobos, 2004). Por lo tanto, los test de autoinforme son, a día de hoy, necesarios e indispensables para poder evaluar los estresores de origen laboral.

No obstante, no existe una herramienta de evaluación de riesgos psicosociales comúnmente aceptada, lo que lleva a tener que prestar especial atención a la elección de la más adecuada para cada organización, pues de ellas dependen gran parte de las decisiones empresariales encaminadas a eliminar o minimizar los riesgos psicosociales. La calidad de los test, junto con

el papel desempeñado por la persona encargada de administrarlos y corregirlos, es fundamental para garantizar unos resultados confiables. Esto hace suponer que una mala utilización de los mismos podría llevar a tomar decisiones erróneas, y así incidir negativamente tanto en la salud y bienestar de los trabajadores como en los aspectos productivos y económicos de la empresa. Cabe recordar que en la última ECVT (2010) publicada, un 49.90% de trabajadores declararon padecer niveles altos o muy altos de estrés, el cual está asociado a multitud de dolencias que merman la salud y el bienestar (véase epígrafe 1.4.1.2), además de provocar grandes depresiones económicas en las empresas e inmensos costes sanitarios (ENWHP, 2011; MATRIX, 2013; ORP-UGT, 2013). Por lo tanto, recae un cometido de gran responsabilidad sobre las personas encargadas de seleccionar las herramientas de evaluación adecuadas.

Para salvar en lo posible esta barrera, los responsables en materia de prevención pueden utilizar como orientación y guía, la revisión efectuada en el epígrafe 2.4. Dicha revisión analiza de forma detallada las principales herramientas de evaluación psicosocial recomendadas por el ITSS (2012), dando a conocer su fundamentación, teórica, sus propiedades psicométricas, las variables que miden y el número de ítems, entre otras características. Tal vez los técnicos de prevención de riesgos más veteranos conozcan en detalle cada una de las herramientas, pero sin duda la revisión efectuada puede resultar de inestimable ayuda a los profesionales noveles, pues podrán escoger de manera objetiva la metodología más adecuada. Con un simple vistazo se puede explorar de modo asequible el rigor científico de cada herramienta y, si se adecúa o no a las necesidades de evaluación de la organización en cuestión. El mero hecho de comparar unos cuestionarios de evaluación con otros favorecería en gran medida la tarea de selección de la herramienta más idónea, sobre todo teniendo en cuenta que dicho análisis ha puesto en entredicho el rigor científico de algunas de las herramientas analizadas. También se observan grandes diferencias entre ellas, desde las que cuentan con pobres propiedades psicométricas y miden cuestiones ajenas a las condiciones laborales, hasta las que gozan de idónea calidad psicométrica y se ajustan a los preceptos de la LPRL. A su vez, puede servir de ayuda a la propia Inspección laboral, para actualizar y/o modificar las recomendaciones de metodologías incluidas en su Guía de actuación sobre riesgos psicosociales (ITSS, 2012)

La disparidad de herramientas de evaluación existentes y las discrepancias entre ellas, están contribuyendo a alimentar lo que Niño (2006) ha denominado “la ceremonia de la confusión”. Esta falta de homogeneidad requiere que se inviertan esfuerzos en legislar o por lo menos, normalizar a través de una norma UNE, un proceso específico de evaluación psicosocial.

3. Deficiencias en los procedimientos de autoevaluación de riesgos psicosociales

La responsabilidad que recae sobre las medidas autoinformadas hace necesario promover su buen uso. De hecho, existen recomendaciones internacionales dirigidas específicamente a solventar la mayor parte de problemas originados por el uso de los test. Son los Estándares para los test psicológicos y educativos (AERA et al., 2014; ITC, 2010). Ofrecen desde directrices para construir y adaptar los test, hasta orientaciones sobre su administración y corrección. En ese sentido, en dichos estándares se incide sobre la importancia de las propiedades psicométricas, así como tener en cuenta los sesgos y garantizar la imparcialidad de los test que sean aplicados a diferentes grupos, estudiando su DIF (ITC, 2010).

Sin embargo, este buen uso de los test se pone en entredicho, según el panorama esbozado por los expertos participantes en el Delphi, quienes desvelan deficiencias importantes en los procedimientos de autoevaluación de riesgos psicosociales. Concretamente, alguna de las cuestiones que referencian son:

- a. Con un 76.92% de acuerdo a favor, los panelistas confirman que muy a menudo se confunde lo que es una evaluación psicosocial con las encuestas de clima laboral. De ésta y otras mezcolanzas de conceptos, se hizo eco Niño (2006), etiquetando la problemática como *la ceremonia de la confusión*. Debido al desconcierto existente no es de extrañar encontrarse, dentro de las baterías de evaluación psicosocial, medidas de salud, de estrés o de bienestar de los trabajadores, entre otras variables ajenas a las que analizan específicamente condiciones de trabajo. Si bien este hecho es común, es del todo irregular, contraviniendo los preceptos de la LPRL y de la ITSS (2012), dando lugar en más de una ocasión a situaciones de intromisión profesional (Niño, 2002).
- b. Las condiciones de administración de los cuestionarios de autoinforme tienen una sustancial importancia. A menudo se descuidan aspectos como el ruido, la ansiedad de los trabajadores, las condiciones térmicas del local destinado a administrar los cuestionarios o la preparación del encuestador entre otras circunstancias, y sin embargo tienen gran influencia en la validez de los resultados de las evaluaciones. De hecho, de todas las proposiciones presentadas a los expertos, las concernientes a la aplicación de los test, han sido las que han contado con mayor consenso. De ahí que un 81.25% de los profesionales y académicos encuestados estuviesen de acuerdo en que un protocolo de administración que contribuiría a minimizar los posibles sesgos de respuesta donde, por ejemplo, los encargados de administrar los test puedan leer las instrucciones pausada y claramente y traten de reducir la ansiedad de los evaluados.
- c. Con un notable 92% de acuerdo, los expertos indican que los instrumentos de evaluación creados desde el entorno académico formulan ítems y recomendaciones de actuación demasiado generales. Esta queja es indicio de la necesidad de colaboración entre profesionales y académicos. Con la experiencia y conocimientos de ambos, aunados, podrían perfeccionarse la formulación de los ítems y las recomendaciones de intervención asociadas a cada factor de riesgo psicosocial detectado.
- d. Otra de las debilidades que afectan a la validez de las evaluaciones psicosociales es la vulnerabilidad de los test de autoinforme a ser falseadas y a los sesgos de respuesta en general (debilidad contrastada por los expertos con un 81.25% de consenso). La deseabilidad social ha sido el sesgo más estudiado por la comunidad científica (Cano-García, et al., 2005); sin embargo, sobre este sesgo en concreto no hubo ningún tipo de consenso. La controversia generada, sumada a una búsqueda infructuosa sobre evidencias de esta vulnerabilidad en medidas específicas que evalúen condiciones psicosociales adversas, justifican la necesidad de mayores evidencias empíricas sirviendo, además, de punto de partida para el segundo estudio recogido en este trabajo.
- e. Conseguir respuestas sinceras está directamente relacionado con la forma de administrar los cuestionarios y con la credibilidad de la persona encargada de administrarlos. Esta conclusión, apoyada por más de un 85% de acuerdo a favor, resalta la importancia de contar con un protocolo que oriente sobre la mejor manera de administrar las

herramientas de evaluación psicosocial. La variabilidad humana es una fuente de error que puede llevar a distorsionar los resultados de un autoinforme, así como a malentender las instrucciones, equivocarse al marcar la respuesta o responder de modo socialmente deseable. Los propios encargados de administrar los test pueden cometer errores fruto de la falta de la debida empatía o al no explicar con claridad el propósito y las instrucciones de las pruebas. Incluso, las condiciones ambientales del local donde se cubre el cuestionario tienen influencia en la calidad de las respuestas (ej., ruido o calor) (R. Hernández et al., 2014; Morris y Maisto, 2005; Muñiz y Bartram, 2007; Muñiz et al., 2005; Nunnally y Bernstein, 1995). Éstas y otras circunstancias tienen que ser controladas para minimizar la varianza de error en los resultados. Sería deseable contar con una guía de buenas prácticas que contemple de forma exhaustiva las fuentes de error y el modo de minimizarlas.

Para contrarrestar algunas de las deficiencias señaladas por los expertos, los mismos, han recomendado seguir una serie de actuaciones, siendo la más destacable la siguiente: *Es fundamental e imprescindible, ampliar el proceso de evaluación de factores psicosociales, ir más allá de un cuestionario, y profundizar con los trabajadores en las condiciones y entorno en el que trabajan. Llevar a cabo sesiones de trabajo en las que los trabajadores puedan participar y hacerles partícipes del proceso. La temática psicosocial debería tener un mayor peso en las organizaciones, puesto que es el origen que no se ve, de otro tipo de problemáticas organizacionales.*

La triangulación metodológica también se encuentra entre las recomendaciones de mejora al fortalecer la percepción individual, al igual que adaptar el cuestionario a la población que se dirige, cuidar las condiciones de administración de los test o facilitar la participación de los trabajadores en la evaluación. Éstas y otras encomiendas derivadas de los resultados del primer estudio empírico, y en consecuencia fruto del consenso de expertos, se han recopilado en el epígrafe 4.5.3 (véase tabla 51). Seguir estas directrices permitirá mejorar la eficacia y eficiencia en estas evaluaciones.

4. Los expertos demandan un protocolo de uso sencillo para evaluar riesgos psicosociales

Un protocolo de uso sencillo contribuiría en gran medida a paliar muchas de las deficiencias de las que adolecen las evaluaciones de riesgos psicosociales, como los sesgos de respuesta. En este sentido, los panelistas confirman la utilidad de contar con un protocolo de administración de cuestionarios de cara a minimizar los posibles sesgos de respuesta. Justamente de este protocolo se hace eco AERA et al. (2014) e ITC (2010) como parte de los estándares de buenas prácticas relacionadas con el uso de los test en investigación. El INSHT (2015) es otra de las organizaciones que recomienda la planificación de actuaciones antes de comenzar la evaluación de los riesgos psicosociales.

Teniendo en cuenta que los errores pueden surgir en cualquiera de las fases de la evaluación de riesgos, sería conveniente contar con un protocolo de actuación lo más exhaustivo posible. Como sugerencia, se recomendaría que dicho protocolo abordase las siguientes cuestiones: (1) cumplir con los estándares de aplicación de los test; (2) describir detallada de las medidas disponibles a nivel nacional, incluyendo fundamentación teórica, propiedades psicométricas, control de sesgos y extensión; (3) tener en cuenta la población a la que se dirige y las

circunstancias de evaluación, por ejemplo, si por motivos económicos la empresa ha tenido que incumplir con sus obligaciones salariales, los trabajadores estarán menos receptivos, y por lo tanto sus respuestas pueden llegar a estar sesgadas; (4) formar y entrenar previamente a la persona encargada de administrar los test; (5) preparar con anterioridad el lugar dedicado a cubrir los cuestionarios, eliminando las posibles distracciones e incomodidades como ruidos o condiciones térmicas; (6) fomentar una relación empática con los encuestados para intentar rebajar los niveles de ansiedad de las personas que van a cubrir los test; (7) leer las instrucciones de forma clara y contestando las posibles dudas que puedan surgir; (8) explicar la utilidad de los cuestionarios antes de repartirlo; o (9) implicar a la dirección durante al menos alguna fase del proceso de evaluación.

5. Existencia del sesgo DS y DIF en AN en los ítems del cuestionario CoPsoQ-ISTAS21

Otra conclusión relevante extraída del primer trabajo de campo es la discrepancia de opiniones acerca de la influencia de los sesgos de respuesta cuando se evalúan las condiciones laborales. Hecho, que sumado a la falta de estudios empíricos que clarifiquen la problemática, ha marcado el camino hacia una segunda investigación. Con esa premisa, se analizó la versión corta del CoPsoQ-ISTAS21 en busca de posibles efectos del sesgo de deseabilidad social, principal amenaza a la validez de los cuestionarios según la literatura. Asimismo, tras la existencia de estudios que ponen en tela de juicio la veracidad de las puntuaciones recabadas por los autoinformes que miden estrés, estresores o salud, debido a la influencia de la variable de afecto negativo, se optó por realizar análisis de diferencias de funcionamiento de los ítems según el rasgo de AN, para asegurar que la escala es equitativa.

Los resultados indican que la deseabilidad social, sí es un sesgo a tener en cuenta en las evaluaciones de riesgos psicosociales. Se ha demostrado que algunos trabajadores tienen cierta tendencia a marcar las opciones de respuesta que mejor los identifiquen como sujetos valiosos dentro de la organización, y/o de su contexto entorno-familia (ej., doble presencia). Por otra parte, el estudio también confirma la importancia de las características individuales de los trabajadores, pues la exposición a riesgos psicosociales está siendo medida de distinta forma entre los sujetos con alta y baja AN. Así lo indica el DIF significativo y uniforme hallado en el 90% de los ítems del cuestionario analizado. El papel distorsionador de la afectividad negativa ha quedado demostrado en la percepción que tienen los trabajadores de las condiciones laborales, dando peso a las teorías que indican que las medidas de estresores laborales son susceptibles de contener varianza explicada por AN (Brief et al., 1988; Watson y Pennebaker, 1989). Sin embargo, ninguna de las dos variables se ha tenido en cuenta en la construcción de los test, lo que implica que de alguna forma tiene que controlarse su efecto para disminuir el factor de confusión debido al AN, como también deben de ser controlada la varianza debida a la DS. Para ello, pueden servir de ejemplo los análisis llevados a cabo en este segundo estudio empírico.

La detección de condiciones laborales adversas está indicando la presencia de posibles problemas de salud. Por imperativo legal, el empresario tiene que eliminar o intentar reducir esas condiciones indeseables y, si la única fuente de información a la hora de tomar decisiones son los cuestionarios de autoinforme, cobra vital importancia corregir, en la medida de lo posible, los errores derivados de estas medidas. De no hacerlo, los resultados podrían llevar a

realizar intervenciones erróneas. Por ejemplo, si suponemos que la muestra utilizada en el presente estudio fuesen trabajadores de la misma empresa, haber tomado en cuenta los datos sin realizar ningún control de DS ni AN, hubiese arrojado resultados inflados -en este caso “falsos negativos”-, pues los sujetos con alto DS y alto AN son los que perciben de peor forma el grado en que están expuestos a condiciones desfavorables para la salud, en comparación con los otros grupos sin DS y con AN bajo. En consecuencia, estos hallazgos se consideran de gran utilidad para la práctica profesional al aportar evidencias sobre la influencia de estos sesgos, además de medidas para su análisis y control.

6. La mayoría de herramientas de evaluación disponibles no se ocupan de la detección de sesgos

La falta de control de sesgos es otra de las conclusiones derivadas del primer estudio. Cabe recordar que no basta con que los test cumplan con las propiedades psicométricas de fiabilidad y validez; además tienen que carecer de sesgos (Ponsoda, 2009). Éstos pueden originarse debido a multitud de circunstancias, como una redacción ambigua de los ítems, ítems demasiado largos, predisposición a contestar de manera aquiescente o socialmente deseable, entre muchos otros. En este sentido, más del 75% de panelistas se ponen de acuerdo al señalar que no es habitual incluir sistemas de control o detección de sesgos en las evaluaciones de factores de riesgo psicosocial, y que casi toda la atención se centra en cumplir con las garantías necesarias de fiabilidad y validez, dejando de lado el control de sesgos. Sabiendo que (1) ningún test está exento de error (F. Jiménez y Sánchez, 2004); (2) los sesgos pueden invalidar toda o parte de la investigación, y (3) existe diversidad de técnicas para su identificación y control, no se entiende que los sesgos sean obviados con tanta asiduidad (Dupuis et al., 2015).

Si la mayoría de instrumentos de evaluación no contemplan ni la detección ni el control de sesgos de respuesta como la DS, y los técnicos de prevención de riesgos tienen escasos conocimientos para poder corregirlos, se produce una incongruencia que puede cuestionar la validez de los resultados obtenidos en los test de autoinforme. Según el humilde juicio de la autora sólo podría solventarse si (1) el ámbito académico tuviese en cuenta medidas de control y corrección de sesgos de respuesta en el diseño de los test de autoinforme destinados a identificar los factores de riesgo psicosocial; (2) requerir los servicios de los profesionales de la prevención de riesgos con formación avanzada en estadística para poder realizar análisis más complejos que permitan interpretar los resultados con mayor profundidad y significación; o (3) ampliar las competencias de los técnicos de prevención de riesgos, mediante la debida formación, para que éstos puedan por sí mismos aplicar, corregir e interpretar los test utilizando las debidas técnicas estadísticas.

7. Los conocimientos de psicometría son importantes para optimizar los resultados de las evoluciones realizadas con autoinformes

Los resultados obtenidos en ambos estudios dejan entrever que las evaluaciones de riesgos laborales no están exentas de error y, siendo conocedores de que estos errores constituyen una gran amenaza a la validez, no basta con administrar un cuestionario e interpretar la información saliente del *software* que corresponda. Hay que intentar minimizar la varianza no deseada. Por lo tanto, entre otros aspectos, es preciso (1) tener en cuenta los sesgos, que pueden darse en cualquier fase del procedimiento de evaluación; e (2) indagar en la percepción de los

trabajadores, aspectos como la motivación, el temor, la resistencia o el estilo de liderazgo, así como sus propias características personales –como la afectividad negativa-, pueden influir en los resultados finales, tal y como se desprende de la presente tesis.

Las conclusiones de ésta han hecho destacar la importancia de la psicometría como ciencia técnica de medición psicológica en la evaluación de riesgos psicosociales. Para muchos técnicos de prevención o expertos en la materia, muchas de las cuestiones abordadas en esta tesis supondrán conceptos novedosos. Sin embargo, son temas tradicionales y muy estudiados en evaluación psicológica y en psicometría. Tal y como indica Meliá (2008) un problema de los instrumentos de evaluación psicosocial es “la escasez de formación especializada por parte de muchos profesionales dedicados a la prevención, en disciplinas tan necesarias como la evaluación psicológica, psicometría, psicología del trabajo o de la seguridad y salud ocupacional” (párr. 4). Por ello, sería necesario trasladar con mayor profundidad y en la medida de lo posible, estas nociones a la esfera de la evaluación psicosocial de origen laboral, tanto al plano teórico como al práctico. Habría que sondear la posibilidad de promover el acceso a estos conocimientos a los técnicos de prevención sin formación previa en psicología.

En síntesis, el conjunto de los resultados de los dos estudios llevados a cabo, permiten aumentar el conocimiento en este ámbito al ofrecer a los profesionales información relevante acerca de las circunstancias que entorpecen las evaluaciones de riesgo psicosocial y este hecho es de tal entidad que por sí solo ya tiene especial relevancia. No obstante, se ha ampliado el recorrido de la investigación al profundizar en el estudio de posibles variables que distorsionan los resultados de las medidas de autoinforme que evalúan condiciones laborales, utilizando para ello variedad de herramientas, muestras y análisis estadísticos que los investigadores y profesionales del sector pueden poner en práctica para cuantificar el grado de distorsión de los resultados y mejorar la validez de los datos obtenidos. Estos logros constituyen un avance hacia una correcta evaluación de riesgos, pues, por una parte, contribuyen a reducir la subjetividad a la hora de determinar los niveles de riesgo, y, por otra parte, ayudan a cuantificar y controlar el grado de sesgo y confusión inherente a la deseabilidad social y al afecto negativo.

A pesar de las evidencias aportadas sobre el impacto de la distorsión, serían bienvenidos nuevos estudios confirmatorios, tanto desde un punto de vista teórico como práctico. La carestía de estudios en este terreno implica que aún queda mucho camino por recorrer, y en ese sentido se espera que esta tesis sirva de motivación para ello.

REFERENCIAS

- Abad, F. J., Garrido, J., Olea, J. y Ponsoda, V. (2006). *Introducción a la Psicometría Teoría Clásica de los Tests y Teoría de la Respuesta al Ítem*. Madrid: Universidad Autónoma de Madrid. Recuperado de https://www.uam.es/personal_pdi/psicologia/cadalso/Docencia/Psicometria/Apuntes/tema1TyP_4.pdf
- Abad, F. J., Olea, J., Ponsoda, V. y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Editorial Síntesis.
- Abekura, H., Tsuboi, M., Okura, T., Kagawa, K., Sadamori, S. y Akagawa, Y. (2011). Association between sleep bruxism and stress sensitivity in an experimental psychological stress task. *Biomedical Research*, 32(6), 395–399.
- Achen, A. C. y Stafford, F. P. (2005). *Data quality of housework hours in the Panel Study of Income Dynamics: Who really does the dishes?* Ann Arbor, Michigan: Panel Study of Income Dynamics. Recuperado de <http://www.psidonline.isr.umich.edu/Publications/Papers/achenproxyreports04.pdf>.
- Acosta, J. M. (2011). *Cómo tratar con personas conflictivas. Guía para reducir el estrés y mejorar las relaciones interpersonales*. Barcelona: Profit Editorial.
- Acosta, H., Torrente, P., Llorens, S. y Salanova, M. (2015). La confianza es pasión: La relación entre confianza organizacional y team work engagement. *Revista Latinoamericana de Psicología Positiva*, 1(2), 7-22.
- Adams, G. (1961). Techniques de minimisation ou d'exploitation des tendances de réponse dans les inventaires structurés d'autoévaluation. [Response tendency minimization or exploitation techniques in structured self-assessment inventories]. *Revue de Psychologie Appliquée*, 11, 233-262.
- Adams, G. R., Ryan, J. H., Hoffman, J. J., Dobson, W. R. y Nielsen, E. C. (1985). Ego identity status, conformity behavior, and personality in late adolescence. *Journal of Personality and Social Psychology*, 47(5), 1091-1104. doi: 10.1037/0022-3514.47.5.1091

- Adriaanse, H., Van Reek, J. V., Zandbelt, L. y Evers, G. (1991). Nurses' smoking worldwide: A review of 73 surveys on nurses' tobacco consumption in 21 countries in the period 1959–1988. *International Journal of Nursing Studies*, 28(4), 361-375.
- Agencia Europea para la Seguridad y la Salud en el Trabajo, EU-OSHA (2005). *Investigación sobre el estrés relacionado con el trabajo*. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas. Recuperado de <https://osha.europa.eu/es/tools-and-publications/publications/reports/203>
- Agencia Europea para la Seguridad y la Salud en el Trabajo, EU-OSHA (2009). *Riesgos nuevos y emergentes para la seguridad y la salud en el trabajo*. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas. Observatorio Europeo de Riesgos. Recuperado de <https://osha.europa.eu/>
- Agencia Europea para la Seguridad y la Salud en el Trabajo, EU-OSHA (2010). *ESENER: Encuesta europea en las empresas sobre riesgos nuevos y emergentes*. Luxemburgo.
- Agencia Europea para la Seguridad y la Salud en el Trabajo, EU-OSHA (2013). *Sondeo de opinión paneuropeo sobre seguridad y salud en el trabajo*. Recuperado de <https://osha.europa.eu/en/safety-health-%20in-figures>
- Agencia Europea para la Seguridad y la Salud del trabajo, EU-OSHA (2014a). *El estrés relacionado con el trabajo: la campaña de la EU-OSHA pide a empresarios y trabajadores que aborden juntos la cuestión*. [Comunicado de prensa] Recuperado de <https://osha.europa.eu/es/press/press-releases/work-related-stress-eu-osha-campaign-calls-on-employers-and-workers-to-tackle-it-together>
- Agencia Europea para la Seguridad y la Salud en el Trabajo, EU-OSHA (2014b). *Calculating the cost of work-related stress and psychosocial risks*. Luxemburgo: Publications Office of the European Union.
- Agencia Europea para la Seguridad y la Salud en el Trabajo, EU-OSHA (2015). *ESENER-2: Segunda Encuesta europea en las empresas sobre riesgos nuevos y emergentes. Primeros resultados*. Luxemburgo. Recuperado de <https://osha.europa.eu/sites/default/files/publications/documents/en/publications/reports/esener-ii-first-findings.pdf/esener-II-first-findings.pdf>

- Agencia Europea para la Seguridad y Salud en el Trabajo. EU-OSHA (2016). *Los riesgos psicosociales y el estrés en el trabajo - Salud y seguridad en el trabajo - EU-OSHA*. Recuperado de <https://osha.europa.eu/es/themes/psychosocial-risks-and-stress>
- Aguayo, M. (2007). Confusión e interacción (1): Qué son, qué suponen y cómo manejarlas en el análisis estratificado. *DocuWeb FABIS*. Recuperado de http://www.fabis.org/html/archivo/s/docuweb/Confu_Inter_2r.pdf
- Ahlberg, J., Lobbzoo, F., Ahlberg, K., Manfredini, D., Hublin, C., Sinisalo, J. y Savolainen, A. (2013). Self-reported bruxism mirrors anxiety and stress in adults. *Medicina Oral, Patología Oral y Cirugía Bucal*, 18(1), E7-E11. doi:10.4317/medoral.18232
- Ahlberg, J., Rantala, M., Savolainen, A., Suvinen, T., Nissinen, M., Sarna, S., ... Kononen, M. (2002). Reported bruxism and stress experience. *Community Dentistry and Oral Epidemiology*, 30, 405–408.
- Akerstedt, T. (2003). Shift work and disturbed sleep/wakefulness. *Occupational Medicine*, 53(2), 89-94.
- Åkerstedt, T. (2006). Psychosocial stress and impaired sleep. *Scandinavian Journal of Work Environmental Health*, 32(6), 493–501.
- al'Absi, M., Carr, S. B. y Bongard, S. (2007). Anger and psychobiological changes during smoking abstinence and in response to acute stress: prediction of smoking relapse. *International Journal of Psychophysiology*, 66, 109-115.
- Albertsen, K., Rafnsdóttir, G. L., Grimsø, A., Tómasson K. y Kauppinen K. (2008). Workhours and worklife balance. *Scandinavian Journal of Work, Environment and Health*, 5, 14–21.
- Aldana, S. G., Sutton, L. D., Jacobson, B. H. y Quirk, M. G. (1996). Relationships between leisure time physical activity and perceived stress. *Perceptual and Motor Skills*, 82, 315–321.
- Aliaga, J. (2006). Psicometría: Test psicométricos, confiabilidad y validez. En A. Quintana y W. Montgomery (Eds.), *Psicología, Tópicos de Actualidad* (pp. 85-108). Lima, Perú: Lima: UNMSM. Recuperado de http://blog.uca.edu.ni/kurbina/files/2011/06/test-psicometrico_confiabilidad-y-validez.pdf

- Ali, A. K. (2005). Using the Delphi technique to search for empirical measures of local planning agency power. *The Qualitative Report*, 10(4), 718-744. Recuperado de <http://nsuworks.nova.edu/tqr/vol10/iss4/5/>
- Allaman, J. D., Joyce, C. S. y Crandall, V. C. (1972). The antecedents of social desirability response tendencies of children and young adults. *Child Development*, 43(4), 1135-1160. doi: 10.1111/1467-8624.ep12114746
- Alles, M. (2008). *Dirección estratégica de recursos humanos: casos: gestión por competencias*. (3ª ed. Rev.). Argentina: Ediciones Granica.
- Almodóvar, A., Pinilla, J., Galiana, M. L. y Hervás, P. (2015). Encuesta Nacional de Gestión de Riesgos Laborales en las Empresas (ESENER 2) Comparación entre la UE-28 y España. *Seguridad y Salud en el Trabajo*, 85, 20-29.
- Alwin, D. F. y Krosnick, J. A. (1991). The reliability of survey attitude measurement. *Sociological Methods and Research*, 20(1), 139-181.
- Amagasa, T., Nakayama, T. y Takahashi, Y. (2005). Karojisatsu in Japan: Characteristics of 22 Cases of Work- Related Suicide. *Journal Occupational Health*, 47, 157–164.
- American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing* (3ª Ed.). Washington, DC: American Educational Research Association.
- American Psychological Association, APA (2010). *Ethical principles of psychologists and code of conduct*. Washington: APA. Recuperado de <http://www.apa.org/ethics/code/>
- Ames, G. M., Cunradi, C., Moore, R. y Duke, M. (2009). The Impact of Occupational Culture on Drinking Behavior of Young Adults in the U.S. Navy, *Journal of Mixed Methods*, 3(2), 120-150.
- Ames, G. M. y Delaney, W. (1992). Minimization of workplace alcohol problems: the supervisor's role. *Alcoholism: Clinical and Experimental Research*, 16, 180-189.
- Ames, G. M., Grube, J. y Moore, R (2000). Social control and workplace drinking norms: a comparison of two organizational cultures. *Journal of Studies on Alcohol*, 61, 203-219.

- Anastasi, A. (1988). *Psychological testing* (6th Ed.). New York: Macmillan.
- Anastasi, A. y Urbina, S. (1998). *Test psicológicos*. México: Pretince Hall.
- Anderson, J. A. (1996). *Communication Theory: Epistemological Foundations*. Nueva York: Guilford.
- Anderson, P., Whitmer, W., Goetzel, R., Ozminkowski, R. J. y Wasserman, J. (2000). The relationship between modifiable health risk and group-level health care expenditures. *American Journal of Health Promotion*, 15(1), 45-52.
- Anguiano-Carrasco, C., Vigil-Colet, A. y Ferrando, P. J. (2013). Controlling social desirability may attenuate faking effects a study with aggression measures. *Psicothema*, 25(2), 164-170.
- Anton, S. D. y Miller, P. M. (2005). Do negative emotions predict alcohol consumption, saturated fat intake and physical activity in older adults? *Behavior Modification*, 29, 677-688.
- Appels, A. (1993). Exhaustion as endpoint of job stress and precursor of disease. En L. Levi y LaFerla (Eds.), *A healthier work environment* (pp. 258-265). Copenhagen: WHO.
- Appels, A., Kop, W. J. y Schouten, E. (2000). The nature of the depressive symptomatology preceding myocardial infarction. *The Journal of Behavioral Medicine*, 26(2), 86-89.
- Argibay, J. C. (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad. *Subjetividad y Procesos Cognitivos*, 8, 15-33.
- Arregui, G., Vallejo, B. y Villarreal, O. (1996). Aplicación de la metodología Delphi para la integración española en la Unión Económica Monetaria. *Investigaciones Europeas de Dirección y Economía de La Empresa*, 2(2), 13-37.
- Artazcoz, L. y Molinero, E. (2004). Evaluación de los factores de riesgo psicosocial combinando metodología cuantitativa y cualitativa. *Archivos de Prevención de Riesgos Laborales*, 7(4), 134-142.
- Ashforth, B. (1994). Petty tyranny in organizations. *Human Relations*, 47, 755-778.

- Astigarraga, E. (2003). *El método delphi*. San Sebastián: Universidad Deusto. Recuperado de http://www.unalmed.edu.co/~poboyca/documentos/documentos1/documentos-Juan%20Diego/Plnaifi_Cuencas_Pregado/Sept_29/Metodo_delphi.pdf
- Aveni, M. y Albani, F. (1992). The burnout syndrome in the helping professions: A study of socio-sanitary staff working with the elder persons in different geriatric services in Lombardy area. En Via, J. y Portella, E., *La sociedad ante el envejecimiento y la minusvalía*. (Vol. I). Barcelona: S.G. Editores.
- Ayidiya, S. A. y McClendon, M. J (1990). Response effects in mail surveys. *Public Opinion Quarterly*, 54, 229-247.
- Bacharach, S. B., Bamberger, P. y Sonnenstuhl, W. J. (2002). Driven to drink: managerial control, work-related risk factors, and employee problem drinking. *The Academy of Management Journal*, 45, 637-658.
- Bachman, J. G. y O'Malley, P. M. (1984). Yea-saying, nay-saying, and going to extremes: Black-white differences in response styles. *Public Opinion Quarterly*, 48, 491-509.
- Bandalos, D. L. y Finney, S. J. (2010). Factor Analysis: Exploratory and Confirmatory. En G. R. Hancock y R. O. Mueller (Eds.), *Reviewer's guide to quantitative methods*. Routledge: New York.
- Banta, T. J. (1961). Social attitudes and response styles. *Educational and Psychological Measurement*, 21, 543-557.
- Baratech, M. (2006). El estrés: epidemia laboral del siglo XXI. *Fomento de la Producción*, 1262, 10-11.
- Bardera, P., Osca, A. y González-Camino, G. (2002). Influencia del estrés y la satisfacción laboral sobre la propensión al abandono, el absentismo y la accidentalidad. *Ansiedad y Estrés*, 8(2-3), 275-284.
- Barker, C., Pistrang, N. y Elliott, R. (2002). *Research Methods in Clinical Psychology: An Introduction for Students and Practitioners*. (2nd ed.). Chichester, England: John Wiley & Sons.

- Barling, J. y Griffiths, A. (2003). History of Occupational Health Psychology. En J. C. Quick y L. E. Tetrick (Eds.), *Handbook of occupational health psychology* (pp. 19-33). Washington, DC: APA.
- Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, 60, 361-370.
- Barreiro, G. (dir.) (2006). *Las enfermedades del trabajo: nuevos riesgos psicosociales y su valoración en el derecho de la protección social*. España. Secretaría de Estado de la Seguridad Social, Ministerio de Trabajo y Asuntos Sociales.
- Bartlett, C. J. y Richard, D. (1967). Social desirability response differences under research, simulated selection, and faking instructional sets. *Personnel Psychology*, 20(3), 281-288.
- Bartolomé, M., Fernández, C., Prieto, S. y Sánchez, V. (2007). Riesgos psicosociales emergentes en el trabajo. *Revista universitaria de ciencias del trabajo*, 8, 341-356. Recuperado de [http://www.ruct.uva.es/pdf/Revista 8/8414.pdf](http://www.ruct.uva.es/pdf/Revista%208/8414.pdf)
- Bass, B. M. (1955). Authoritarianism or acquiescence? *Journal of Abnormal and Social Psychology*, 51, 616-623.
- Bass, B. M. (1956). Development and evaluation of a scale for measuring social acquiescence. *Journal of Abnormal and Social Psychology*, 53, 296-299.
- Bass, B. M. (1990). *Handbook of leadership: A survey of theory and research*. New York: Free Press.
- Batchelor, J. H., Miao, C. y McDaniel, M. A. (2013, April). Extreme response style: A meta-analysis. Presentado en la 28th Annual Conference of the Society for Industrial and Organizational Psychology. Houston. Recuperado de http://www.people.vcu.edu/~mamcdani/CV_and_Research_Publications.htm
- Baumgartner, H. y Steenkamp, J.-B. E. M. (2001). Response styles in marketing research: a cross-national investigation, *Journal of Marketing Research*, 38(2) 143-156.

- Beer, A., Watson, D. y McDade-Montez, E. (2013). Self-other agreement and assumed similarity in neuroticism, extraversion, and trait affect: distinguishing the effects of form and content. *Assessment*, 20, 723–37. doi: 10.1177/1073191113500521
- Behar, R. y Grima, P. (2012). *55 Respuestas A Dudas Típicas En Estadística*. 2ª Ed. Editorial Díaz de Santos, Madrid.
- Behar, R. y Valdés, C. (2009). Estrés y trastornos de la conducta alimentaria. *Revista Chilena de Neuro-Psiquiatría*, 47(3), 178-189.
- Belkic, K., Landsbergis, P. A., Schnall, P. L. y Baker, D. (2004). Is job strain a major source of cardiovascular disease risk? *Scandinavian Journal of Work and Environmental Health*, 30(2), 85-128.
- Benach, J. y Muntaner, C. (2007). Precarious employment and health: developing a research agenda. *Journal of epidemiology and community health*, 61, 276–7.
- Benavides, F., Gimeno, D., Benach, J., Martínez, J. M., Jarque, S. y Berra, A. (2002). Descripción de los factores de riesgo psicosocial en cuatro empresas. *Gaceta Sanitaria*, 16, 222-229.
- Bennett, J. B. y Lehman, W. E. K. (1998). Workplace drinking climate, stress, and problem indicators: assessing the influence of teamwork (group cohesion). *Journal of Studies on Alcohol and Drugs*, 10, 135-143. doi: 10.15288/jsa.1998.59.608
- Bennett, J. B., Cook, R. F. y Pelletier, K. R. (2003). Toward an Integrated Framework for Comprehensive Organizational Wellness: Concepts, Practices, and Research in Workplace Health Promotion. En J. C. Quick y L. E. Tetrick (Eds.), *Handbook of occupational health psychology* (pp. 69-95). Washington, DC: APA.
- Bennett, J. B., Paterson, C. R., Reynolds, G. S., Wiitala, W. L. y Lehman, W. E. K. (2004). Team awareness, problem drinking, and drinking climate: workplace social health promotion in a policy context. *The American Journal of Health Promotion*, 19, 103-113.
- Bentler, P. M. (1977). Factor simplicity index and transformations. *Psychometrika*, 42(2), 77-295.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 109, 238-246.

- Bernreuter, R. G. (1933). Validity of the personality inventory. *Personality Journal*, 11, 383-386.
- Berthelsen, H., Hakanen, J. Kristensen, T. S., Lönnblad, A. y Westerlund, H. (2016). A Qualitative Study on the Content Validity of the Social Capital Scales in the Copenhagen Psychosocial Questionnaire (COPSOQ II). *Scandinavian Journal of Work and Organizational Psychology* 1(1), A5,1-13. doi: 10.16993/sjwop.5
- Bestraten, M., Bernal, F., Castillo, M., Cejalvo, A., Hernández, A., Luna, ...Turmo, E. (2000). *Evaluación de las condiciones de trabajo en pequeñas y medianas empresas* (5ª ed.). Barcelona: INSHT.
- Bies, R. J. y Tripp, T. M. (2005). The study of revenge in the workplace: Conceptual, ideological, and empirical issues. En P. E. Spector y S. Fox (Eds.), *Counter-productive work behavior: Investigation of actors and targets* (pp. 65–105). Washington, DC: APA Books.
- Bilkei-Gorzo, A., Rácz, I., Michel, K., Darvas, M., Maldonado, R. y Zimmer, A. (2007). A common genetic predisposition to stress sensitivity and stress- induced nicotine craving. *Biological Psychiatry*, 63, 164-171. doi:10.1016/j.biopsych.2007.02.010.
- Billiet, J. B. y Davidov, E. (2008). Testing the Stability of an Acquiescence Style Factor Behind Two Interrelated Substantive Variables in a Panel Design. *Sociological Methods Research*, 3(4), 542-562. doi: 10.1177/0049124107313901
- Billiet, J. B. y McClendon, M. J. (1998). On the identification of acquiescence in balanced set items using structural models. En A. Ferligoj. (ed.), *Advances in Methodology, Data Analysis, and Statistics. Metodoloski zvezki* (Vol. 14, pp. 129–150). Ljubljana: FDV
- Bishop, G. D., Enkelmann, H. C., Tong, E. M., Why, Y. P., Diong, S. M., Ang, J. y Khader, M. (2003). Job demands, decisional control, and cardiovascular responses. *Journal of Occupational Health Psychology*, 8(2), 146–156.
- Björkqvist, K., Österman, K. y Hjelt-Bäck, M. (1994). Aggression among university employees. *Aggressive Behavior*, 20, 173-184.
- Blackmore, E. R., Stansfeld, S. A., Weller, I., Munce, S., Zagorski, B. M. y Stewart, D. E. (2007). Major Depression Episodes and Work stress: Results From a National Population Survey. *American Journal of Public Health*, 97, 11, 2088-2093.

- Blanch, J. M. (2003). Trabajar en la sociedad informacional. En J. M. Blanch, M. J. Espuny, C. Gala y A. Martín (Eds.), *Teoría de las relaciones laborales: Desafíos*. Barcelona: Editorial UOC.
- Blanch, J. M., Sahagún, M. y Cervantes, G. (2010). Estructura Factorial del Cuestionario de Condiciones de Trabajo. *Revista de Psicología del Trabajo y de las Organizaciones*, 26(3), 175–189. doi:10.5093/tr2010v26n3a2
- Blanch, J. M., Sahagún, M. A., Cantera, L. y Cervantes, G. (2010). Cuestionario de Bienestar Laboral General: Estructura y Propiedades Psicométricas. *Revista de Psicología del Trabajo y las Organizaciones*, 26(2), 157-170.
- Blanco, A., España, E., Franco, A. J. y González, F. J. (2014). Aspectos metodológicos en un estudio Delphi sobre la competencia científica deseable para la ciudadanía. En Heras, M.A, Lorca, A., Vázquez, B., Wamba, A. y Jiménez, R. (coords). *Investigación y transferencia para una educación en ciencias. Un reto emocionante* (pp. 176-183). Huelva: Publicaciones de la Universidad de Huelva
- Blasco, J. E., López, A. y Mengual, S. (2010). Validación mediante el método Delphi de un cuestionario para conocer las experiencias e interés hacia las actividades acuáticas con especial atención al windsurf. *Ágora para la educación física y le deporte*, 12(1), 75-94.
- Boada, J., De Diego, R. y Vigil, A. (2003). Mobbing: Análisis de las propiedades psicométricas y estructura factorial de cuatro escalas (MOBB-90; MobbCF-21; MobbCG-15 y MobbCS-28). *Revista de Encuentros de Psicología Social*, 1, 26-31.
- Boix, P., Gil, J. M. y Rodrigo, F. (2008). Prioridades estratégicas para la mejora del sistema de prevención de riesgos laborales en España: un estudio Delphi entre profesionales de la salud y seguridad en el trabajo. *Archivos de Prevención de Riesgos Laborales*, 11(1), 20-26.
- Bonde, J. P. (2008). Psychosocial factors at work and risk of depression: a systematic review of the epidemiological evidence. *Journal of Occupational and Environmental Medicine*, 65 (7), 438-445.
- Bongers, P. M. (2005). Psychosocial work characteristics in relation to neck and upper limb symptoms. *Pain*, 114, 47-53.

- Bongers, P. M., de Winter, C. R., Kompier, M. A. J. y Hildebrandt, V. H. (1993). Psychosocial factors at work and musculoskeletal disease. *Scandinavian Journal of Work, Environment and Health*, 19(5), 297-312. doi:10.5271/sjweh.1470
- Borkenau, P. y Ostendorf, F. (1989). Descriptive consistency and social desirability in self- and peer reports. *European Journal of Personality*, 3, 31-45.
- Bosques, F. J. y Gómez, D. (2009.) El sesgo de publicación es un problema científico frecuente que tiene consecuencias éticas adversas. *Medicina universitaria*, 11(42), 1-2.
- Bourdieu, P. (1986). The forms of capital. En J. Richardson (Ed.) *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258) New York, Greenwood.
- Bourgeois, J., Pugmire, L., Stevenson, K., Swanson, N. y Swanson, B. (2006). *The Delphi Method: A Qualitative Means to a Better Future*. Recuperado <http://staff.estem-uc.edu.au/taipham/files/2013/01/Delphimethod.pdf>
- Bray, R. M., Fairbank, J. A. y Marsden, M. E. (1999). Stress and substance use among military women and men. *American Journal of Drug and Alcohol Use*, 25(2), 239-256. Recuperado de https://www.rti.org/pubs/bray-stresssubst_abuse_military.pdf
- Brengelmann, J. C. (1986). Estrés, superación y calidad de vida en personas sanas y enfermas. *Evaluación psicológica*, 2, 47-78.
- Brewer, E. W. (2007). Delphi technique. En N. J. Salkind & K. Rasmussen (Eds.), *Encyclopedia of Measurement Statistics* (Vol. 1, pp. 240-246). Sage, Thousand Oaks: Sage.
- Brief, A. P., Burke, M. J., George, J. M., Robinson, B. S. y Webster, J. (1988). Should negative affectivity remain an unmeasured variable in the study of job stress? *Journal of Applied Psychology*, 73, 193-198.
- Brill, J., Bishop, M. y Walker, A. (2006). An investigation into the competencies required of an effective project manager: A Web-based Delphi study. *Educational Technology Research & Development*, 54(2), 115-140.

- Brisson, C., Larocque, B., Moisin, J., Vexina, M. y Dagenis, G. R. (2000). Psychosocial factors at work, smoking, sedentary behavior and bodymass index: A prevalence study among 6995 white-collar workers. *Journal of Occupational and Environmental Medicine*, 42, 40–46.
- Broadbent, D. E., Cooper, P. F., FitzGerald, P. y Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ). *British Journal of Clinical Psychology*, 21, 1-16.
- Brockhoff, K. (1975). The performance of forecasting groups in computer dialog and face-to-face discussion. En H. A. Linstone & M. Turoff (Eds.), *The Delphi method: Techniques and applications* (pp. 291-321). Reading, MA: Addison-Wesley
- Broeck, J. v. d. y Brestoff, J. R. (2013). *Epidemiology: Principles and practical guidelines*. Dordrecht, New York: Springer. doi: 10.1007/978-94-007-5989-3
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Brunborg, G. S. (2008). Core self-evaluations. A predictor variable for job stress. *European Psychologist*, 13(2), 96-102.
- Brunner, E. J., Chandola, T. y Marmot, M. G. (2007). Prospective Effect of Job Strain on General and Central Obesity in the Whitehall II Study. *American Journal of Epidemiology*, 165(7), 828–837.
- Bucci, L. M. (2012). Contrastación entre los modelos de estudio del estrés como soporte para la evaluación de los riesgos psicosociales en el trabajo. *Redip. Unexpo. Vrb. Venezuela*, 2(1), 21–38.
- Buchanan, T. y Smith, J. L. (1999). Using the Internet for psychological research: Personality testing on the World Wide Web. *British Journal of Psychology*, 90, 125-144.
- Buela-Casal, G. y Sierra, J. C. (1997). *Manual de Evaluación Psicológica: Fundamentos, técnicas y aplicaciones*. Madrid: Siglo XX.
- Bueno, R. (2014). Ciencia, ideología e investigación social: Comentarios sobre un artículo de Chavarría (2011). *Actualidades en Psicología*, 28, 67-80.

- Bunge, M. (1973). *La ciencia, su método y filosofía*. Buenos Aires: Siglo XX.
- Bunker, S. J., Colquhoun, D. M., Esler, M. D., Hickie, I. B., Hunt, D., Jelinek., ... Tonkin, A. M. (2003). "Stress" and coronary heart disease: Psychosocial risk factors. *Medical Journal of Australia*, 178(6), 272–276.
- Burgard, S., Brand, J. y House, J. (2009). Perceived job insecurity and worker health in the United States. *Social Science y Medicine*, 69, 777–785.
- Burke, R. J. (2001). Estilos de Afrontamiento. En S. L. Sauter, L. R. Murphy, J.J. Hurrell y L. Levi (Dirs.). *Factores psicosociales y de organización. Enciclopedia de Salud y Seguridad en el Trabajo* (cap. 34, pp. 54-55). Madrid: Organización Internacional del Trabajo.
- Caballero, C. C., Hederich, C. y Palacio, J. E. (2010). El burnout académico: delimitación del síndrome y factores asociados con su aparición. *Revista Latinoamericana de Psicología*, 42(1), 131-146.
- Cabañeros, J., García Cueto, E. y Lozano, L. M. (2003). El efecto de la deseabilidad social en preguntas sobre temas comprometidos. Ponencia presentada en la *IX Conferencia Española de Biometría*. La Coruña, mayo, 2003.
- Cabero, J. y Barroso, J. (2013). La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta. *Bordón*, 65(2), 25-38.
- Cabero, J. y E. Infante, A. (2014). Empleo del método Delphi y su empleo en la investigación, en comunicación y educación. *Revista Edutec*, 48, 1-16.
- Cabero, J. y Llorente, M. C. (2013). La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información (TIC). En Eduweb. *Revista de Tecnología de Información y Comunicación en Educación*, 7(2)11-22. Recuperado de <http://servicio.bc.uc.edu.ve/educacion/eduweb/v7n2/art01.pdf>
- Caicoya, M. (2004). Dilemas en la evaluación de riesgos psicosociales. *Archivos de prevención de riesgos laborales*, 7(3), 109-118.
- Calabuig, F. y Crespo, J. (2009). Uso del método Delphi para la elaboración de una medida de la calidad percibida de los espectadores de eventos deportivos Retos. *Nuevas Tendencias en Educación Física, Deporte y Recreación*, 15, 21–25.

- Calderón, C. y Ximénez, C. (2014). Análisis factorial de ítems de respuesta forzada: una revisión y un ejemplo. *Revista Latinoamericana de Psicología*, 45, 24-34 doi:10.1016/S0120-0534(14)70003-2
- Callejón-Ferre, A. J., Montoya-García, M. E., Pérez-Alonso, J. y Rojas-Sola, J. I. (2015). The psychosocial risks of farm workers in south-east Spain. *Safety Science*, 78, 77-90.
- Camisón, C., Camisón, C., Fabra, E., Florés, B. y Puig, A. (2009) ¿Hacia dónde se dirige la función de calidad? La visión de expertos de un estudio Delphi. *Revista Europea de Dirección y Economía de la Empresa*, 18(2), 13-38.
- Cameron, H. L. y Perdue, M. H. (2005). Stress impairs murine intestinal barrier function: improvement by glucagon-like peptide-2. *J. Journal of Pharmacology and Experimental Therapeutics*, 314(1), 214-220.
- Canavos, G. C. (2003). *Probabilidad y estadística. Aplicaciones y métodos*. México: McGraw-Hill.
- Cano-García, F. J., García-Martínez, J., Rodríguez-Franco, L. y Antuña-Bellerín, M. A. (2005). *Introducción a la Psicología de la personalidad aplicada a las ciencias de la Educación: Manual Teórico*. Sevilla: MAD.
- Cano-Vindel, A., Iruarrizaga, I., Ramos, J., Dongil, E., Wood, C. y Salguero, J. M. (2011). Estrés, ansiedad, depresión y vulnerabilidad al accidente. En F. Tortosa-Gil y C. Civera-Mollá (Eds.), *Siniestralidad Vial. Perfiles de Riesgo y Prevención Laboral* (pp. 198-211). Valencia: Alfa-Delta Editores.
- Cano-Vindel, A. y Miguel-Tobal, J. J. (1999). Valoración, afrontamiento y ansiedad. *Ansiedad y Estrés*, 5, 129-143.
- Cannon, W. B. (1932). *The wisdom of the body*. New York: Norton and Co.
- Cantril, H. (1946). The intensity of an attitude. *Journal of Abnormal and Social Psychology*, 41, 129-135.
- Carayon, P., Haims, M. C. y Yang, C. L. (2001). Psychosocial work factors and work organization. En W. Karwowski (Ed.), *The International Encyclopedia of Ergonomics and Human Factors* (pp. 111-121). London: Taylor y Francis.

- Carbone, E. G. y Cigrang, J. A. (2001). Job satisfaction, occupational stress, and personality characteristics of Air Force military training instructors. *Military Medicine*, 166(9), 800-802.
- Carlin, M. y Garcés de los Fayos, E. J. (2010). El síndrome de burnout: Evolución histórica desde el contexto laboral al ámbito deportivo. *Anales de Psicología*, 26(1), 169–180.
- Carrasco, A. M., de la Corte, C. M. y León, J. M. (2010). Engagement: Un recurso para optimizar la salud psicosocial en las organizaciones y prevenir el burnout y estrés laboral. *Revista Digital de Seguridad y Salud en el trabajo*, 1,1-22.
- Carrasquer, P. (2002). ¿En los límites de la modernidad? Trabajo y empleo femenino precario en España. *Sistema* 167, 77-99.
- Carreño, S., Medina-Mora, M. E, Martínez, N. A., Juárez, F. y Vázquez, L. (2006). Características organizacionales, estrés y consumo de alcohol en trabajadores de una empresa textil mexicana. *Salud Mental*, 29(4), 63-70.
- Castillo, C. G., Díaz, M. H., Fariñas, A. G. y Delgado, Z. S. (2009). Propuesta de competencias laborales para médicos que brindan servicios de atención médica en hoteles cubanos. *Revista Cubana de Educación Médica Superior*, 23(3), 15–26.
- Castro, L. I., Pedraza, P., Ortiz, R., Bermúdez I. B., Reyes, I., Salas-rojas, S. G. y Céspedes, L. M. (2014). Farmacéuticas. *Gestión de la calidad en Unidades de Farmacovigilancia*, 57–77.
- Carvalho, A. L. D. A., Cury, A. A. D. B. y García, R. C. M. R. (2008). Prevalence of bruxism and emotional stress and the association between them in Brazilian police officers. *Brazilian Oral Research*, 22(1), 31–35.
- Cattell, R. B., Eber, H. W. y Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF)*. Champaign, Illinois: Institute for Personality for y and Ability Testing (IPAT).
- Chaffin, W.W. y Talley, W.K. (1980). Individual Stability in Delphi studies. *Technological Forecasting & Social Change*, 16, 67–73.

- Chami-Castaldi, E., Reynolds, N. y Wallace, J. (2008). Individualised rating-scale procedure: A means of reducing response style contamination in survey data. *Electronic Journal of Business Research Methods*, 6(1), 9–20.
- Chandola, T., Britton, A., Brunner E., Hemingway, H., Malik, M., Kumari, M, ... Marmot, M. (2008). Work stress and coronary heart disease: what are the mechanisms? *European Heart Journal*, 29(5), 640-648. doi:10.1093/eurheartj/ehm584.
- Charro, B. (2012). El consumo de sustancias psicoactivas como riesgo psicosocial en al ámbito laboral español. una revisión. *International Psychology, Practice and Research*, 3, 1-13. Recuperado de http://fiuc.org/es/piper/fourth_trimester_2012
- Chavarría, M. C. (2011). La dicotomía cuantitativa/cualitativo, falsos dilemas en la investigación social. *Actualidades en psicología*, 25(112), 1-35. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0258-64442011000100001&lng=pt&tlng=es
- Cherns, A. (1976). The principles of sociotechnical design. *Human Relations*, 29(8), 783-792.
- Cheung, M. y Chan, W. (2002). Reducing uniform response bias with ipsative measurement in multiple-group confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(1), 55-77. doi: 10.1207/S15328007SEM0901_4
- Cheung, G. W. y Rensvold, R. B. (1999). Testing factorial invariance across groups: a reconceptualization and proposed new method. *Journal of Management*, 25, 1- 27.
- Choi, B., Granero, R. y Pak, A. (2010). Catálogo de sesgos o errores en cuestionarios sobre salud. *Revista Costarricense de Salud Pública*, 2(19), 106–118.
- Chou, C.P. y Bentler, P. M. (1995). Estimates and tests in structural equation modeling. En R. H. Hoyle (Ed.), *Structural equation modeling Concepts issues and applications* (pp. 37-55). Thousand Oaks, CA: Sage Publications.
- Christianson, S. A. (1992). Emotional stress and eyewitness memory: a critical review. *Psychological Bulletin*, 112, 284–309.

- Cifre, E., Salanova, M. y Franco, J. (2011). Riesgos psicosociales de hombres y mujeres en el trabajo: ¿Una cuestión de diferencias? *Revista Gestión Práctica de Riesgos Laborales*, 82, 28–37.
- Cifre, E., Salanova, M. y Rodríguez, A. (2009). Aplicación del Modelo RED en el proceso de investigación-acción en una esmaltera. En S. Vega (Dir.). *Experiencias en Intervención Psicosocial. Más allá de la evaluación del riesgo*. Barcelona: Instituto Nacional de Seguridad e Higiene en el Trabajo.
- Clark, L. y Watson, D. (1988). Mood and the Mundane: Relationship between Daily Life Events and Self-Reported Mood. *Journal of Personality and Social Psychology*, 54, 296–308.
- Clarke, M. y Oxman, A. D. (Eds.). (2003). *Cochrane reviewers' handbook*. The Cochrane Library (Issue 2). Oxford: Update Software.
- Cliff, N. (1979). La psicometría. En B.B. Wolman (ed.) *Manual de Psicología general. Vol. 1: Historia, teoría y método* (pp. 166-220). Barcelona: Martínez Roca (Orig. 1973).
- Climent, V., Melián A. y Sanchis, J. R. (2014). El método Delphi como técnica de diagnóstico estratégico. Estudio empírico aplicado a las empresas de inserción en España. *Revista Europea de Dirección y Economía de la Empresa*, 23(2), 72–81. doi:10.1016/j.redee.2013.06.002
- Cohen, B. E., Edmondson, D. y Kronish, I. M. (2015). State of the art review: Depression, stress, anxiety, and cardiovascular disease. *American Journal of Hypertension*, 28(11):1295-302. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/25911639>
- Cohen, R. J. y Swerdlik, M. E. (2006). *Pruebas y Evaluación Psicológicas*. McGraw-Hill. México D.C.
- Cohen, S. (1980). Aftereffectsof stress on human performance and social behavior: a review of research and theory. *Psychological Bulletin*, 88, 81-108.
- Cohen, S., Hamrick, N. y Rodríguez, M. S. (2002). Reactivity and vulnerability to Medicine stress-associated risk for upper respiratory illness. *Psychosomatic Medicine*, 64, 302–310.

- Colell, E., Sánchez-Niubò, A., Domingo-Salvany, A., Delclós, J. y Benavides, F. G. (2014). Prevalencia de consumo de hipnosedantes en población ocupada y factores de estrés laboral asociados. *Gaceta Sanitaria*, 28(5), 369-375. doi:10.1016/j.gaceta.2014.04.009
- Colell, E., Sánchez-Niubò, A., Benavides, F. G., Delclos, G. L. y Domingo-Salvany, A. (2014). Work-related stress factors associated with problem drinking: A study of the Spanish working population. *American Journal of Industrial Medicine*, 57(7), 837-846 doi: 10.1002/ajim.22333
- Coleman, J. S. (1958). Snowball sampling—Problems and techniques of chain referral sampling. *Human Organization*, 17, 28-36.
- Coleman, J. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94 (Supplement), 95-120.
- Collazo, A. A. (2005). Translation of the Marlowe-Crowne social desirability scale into an equivalent Spanish version. *Educational and Psychological Measurement*, 65, 780-806.
- Colligan, M. J., Smith, M. J. y Hurrell, J. J. (1977). Occupational incidence rate of mental health disorders. *Journal of Human Stress*, 34-39.
- Collins, S.M., Barbara. G. y Vallance, B. (1999). Stress, inflammation and the irritable bowel syndrome. *Canadian Journal of Gastroenterology and Hepatology*, 13(Suppl A):47A–49A.
- Collins, S.M., Karasek, R.A. y Costas. K. (2005). Job strain and autonomic indices of cardiovascular disease risk. *American Journal of Industrial Medicine*, 48(3), 182-193.
- Comisión Europea (2017). *Informe sobre España 2017, con un examen exhaustivo relativo a la prevención y la corrección de los desequilibrios macroeconómicos*. Recuperado de <https://ec.europa.eu/info/sites/info/files/2017-european-semester-country-report-spain-es.pdf>
- Comisión Internacional de Tests, ITC. (2010). Directrices internacionales para el uso de los tests. [Traducción de la Comisión de Tests del Colegio Oficial de Psicólogos de España]. *INFOCOP*, 77, 21-32. Recuperado de http://www.infocop.es/view_article.asp?id=5084

- Committee of Senior Labour Inspectors, SLIC. (2012). *Psychosocial risk assessments – SLIC Inspection Campaign 2012: Final report. SLIC*. Recuperado de [http://www.av.se/dokument/inenglish/European Work/Slic 2012/Country reports/Country_report II ES.pdf](http://www.av.se/dokument/inenglish/European_Work/Slic_2012/Country_reports/Country_report_II_ES.pdf).
- Confederación española de organizaciones empresariales, CEOE. (2013). *Análisis de la percepción empresarial en materia de identificación y evaluación de Riesgos Psicosociales: apunte sobre criterios y necesidades metodológicas*. Madrid: CEOE. Recuperado de http://contenidos.ceoe.es/resources/image/estudio_riesgos_piscosociales_ceoe.pdf
- Connolly, J. J. y Viswesvaran, C. (2000). The role of affectivity in job satisfaction: A meta-analysis. *Personality and Individual Differences*, 29, 265-281.
- Consentino, A. C. y Castro-Solano, A. (2008). Adaptación y validación argentina de la Marlowe-Crowne Social Desirability Scale. *Interdisciplinaria*, 25(2), 197-216.
- Converse, P. D., Oswald, F. L., Imus, A., Hedricks, C., Roy, R. y Butera, H. (2008). Comparing personality test formats and warnings: Effects on criterion-related validity and test-taker reactions. *International Journal of Selection and Assessment*, 16, 155-169.
- Cook, J. y Wall, T. (1980). New Work Attitude Measures of Trust, Organizational Commitment and Personal Need Nonfulfillment. *Journal of Occupational Psychology*, 53(1), 39–52.
- Cooper, C.L. (1986). *Estrés y Cancer*. Madrid: Martínez Santos.
- Cooper, C.L. (1993). Identifying Workplace Stress: Costs, Benefits and the Way Forward. *European Conference on Stress at Work. A Call for Action: Proceedings*.
- Cooper, C.L., S. J. Sloan y S. Williams (1988). *The Occupational Stress Indicator (OSI)*, Windsor: NFER Nelson.
- Cooper, C. L. y Smith, M. J. (1986). *Job Stress and Blue Collar Work*. Chichester, UK: Wiley and Sons.

- Cooper, C.L. y Smith, M. (2001). Estilo de dirección. En S.L. Sauter, L.R. Murphy, J.J. Hurrell y L. Levi (Dirs.), *Factores psicosociales y de organización. Enciclopedia de Salud y Seguridad en el Trabajo* (cap.34, pp.39-40) Madrid: Organización Internacional del Trabajo.
- Corbetta, P. (2003). *Metodología y técnicas de investigación social*. Madrid: McGrawHill.
- Corcoran, K. J. (1995). Measuring burnout: An updated reliability and convergent validity study. En R. Crandall y P.L. Perrewe (Eds.), *Occupational stress: A handbook* (pp. 263-268). Philadelphia, PA: Taylor & Francis.
- Cortada de Kohan, N. (1999). *Teorías psicométricas y construcción de tests*. Buenos Aires: Lugar Editorial.
- Cortada de Kohan, N. (2001). Constructos psicológicos y variables. *Psicología y Pedagogía*, 2(2). Recuperado de <http://psico.usal.edu.ar/psico/constructos-psicologicos-variables>
- Cortés, I., Artazcoz, L., Igartua, A., Juanola, E., Ladona, C., Llonch, A., ... Molinero, E. (2010). *Guía de buena praxis para la vigilancia de la salud mental relacionada con los factores de riesgo psicosocial*. Generalitat de Catalunya. Departament de Salut. Direcció General de Salut Pública. Barcelona.
- Cortés J. M. (2012). *Técnicas de prevención de riesgos laborales*. (10ª ed.). Madrid: Tebar.
- Cortés, J. M., Pellicer, E. y Catalá, J. (2012). Integration of occupational risk prevention courses in engineering degrees: Delphi study. *Journal of Professional Issues in Engineering Education and Practice*. 138(1), 31-36. doi:10.1061/(ASCE)EI.1943-5541.0000076
- Cortina, L. M., Magley, V. J., Williams, J. H. y Langhout, R. D. (2001). Incivility in the workplace: Incidence and impact. *Journal of Occupational Health Psychology*, 6, 64-80.
- Cosentino, A. C. y Castro, A. (2008). Adaptación y validación argentina de la Marlowe-Crowne Social Desirability Scale. *Interdisciplinaria*, 25(2), 197-216.
- Costa, G., Ghirlanda, G., Minors, D.S. y Waterhouse, J. M. (1993). Effects of bright light on tolerance to night work. *Scandinavian Journal of Work Environment and Health*, 29,414-420.

- Costa, P. T. y McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social*, 38, 668 – 678.
- Couch, A. y Keniston, K. (1960). Yeasayers and naysayers: Agreeing response set as a personality variable. *The Journal of Abnormal and Social Psychology*, 60(2), 151-174.
- Cousins R., Mackay C. J., Clarke S. D., Kelly C., Kelly P. J. y McCaig R. H (2004). Management Standards and work-related stress in the UK: practical development. *Work & Stress*, 18(2), 113–136. Recuperado de <http://www.hse.gov.uk/stress/techpart2.pdf>
- Cox, T. y Ferguson, E. (1994). Measurement of the subjective work environment. *Work and Stress*, 8(2), 98-109.
- Cox, T. y Griffiths, A. J. (1996). The assessment of psychosocial hazards at work. En M.J. Schabracq, J. A. M. Winnubst y C. L. Cooper (Eds.), *Handbook of Work and Health Psychology* (pp. 127-146). Chichester: Wiley and Sons.
- Cox, T., Griffiths, A. y Leka, S. (2005). Work organization and work related stress. En K. Gardiner and J. M. Harrington (Eds.). *Occupational Hygiene*. Oxford: Blackwell.
- Cox, T., Griffiths, A. y Randall, R. A. (2003) Risk Management Approach to the Prevention of Work Stress. En: Marc J. Schabracq, Jacques A. M. Winnubst, Cary L. Cooper. *The Handbook of Work and Health Psychology* (2nd ed., pp. 191). Chippenham (England): John Wiley & Sons.
- Cox, T., Griffiths, A. y Rial-Gonzalez, E. (2000). *Research on work-related stress*. Luxembourg: Office for Official Publications of the European Communities: European Agency for Safety and Health at Work.
- Cox, T. y Mackay, C. J. (1981). A transactional approach to occupational stress. En J. Corlett y J. Richardson (Eds.), *Stress, productivity and work desing*. Chichester, John Wiley & Sons.
- Crocker, L. y Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Cronbach, L. J. (1946). Response set and test validity. *Educational and Psychological Measurement*, 6, 475-494.

- Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement*, 10, 3-31.
- Cronbach, L.J. (1971). Test validation. En R. L. Thorndike (Ed.), *Educational measurement*. (2nd ed., pp.443-507). Washington, DC: American Council on Education.
- Cronbach, L.J. (1976). Equity in selection: Where psychometrics and political philosophy meet. *Journal of Educational Measurement*, 13(1), 31-41.
- Cronbach, L. J. y Meehl, P. E. (1995). Construct Validity in Psychological Tests. *Psychological Bulletin*, 52, 281-302.
- Crowne, D. P. y Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Crowne, D. P. y Marlowe, D. (1964). *The approval motive: studies in evaluative dependence*. New York: Wiley.
- Cruz, J., de La Flor, M. L. y Álvarez del Cuvillo, A. (2007). Concepto y tratamiento jurídico de los riesgos psicosociales. En AA.VV. (coord. Cruz Villalón, J.) *Análisis de la regulación jurídico-laboral de los riesgos psicosociales. Investigación FIPROS* (cap. I, páginas no numeradas). Madrid. Recuperado de <http://www.seg-social.es/prdi00/groups/public/documents/binario/100617.pdf>
- Cuenca, R. (2002). Los riesgos psicosociales y su prevención: mobbing, estrés y otros problemas. Ponencia presentada en la *Jornada técnica de actualización de los riesgos psicosociales y su prevención: mobbing, estrés y otros problemas*. 10 de diciembre. Madrid. Recuperado de <http://www.carm.es/ctra/cendoc/doc-pdf/web/web-0009.pdf>
- Cunradi, C., Todd, M. Duke, M. y Ames, G. (2009). Problem Drinking, Unemployment and Intimate Partner Violence among a Sample of Construction Workers and their Partners, *Jour of Family Violence*, 24, 63-74.
- Curran, P. J., West, S. G. y Finch, J. F. (1996). The robustness of test statistics to non normality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29.

- Dabos G. E. y Rousseau, D. M. (2004). Mutuality and reciprocity in the psychological contracts of employees and employers. *Journal of Applied Psychology*, 89(1), 52-72.
- Dajani, J. S., Sincoff, M. Z. y Talley, W. K. (1979). Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change*, 13(1), 83-90.
- Dajani, J. S., Sincoff, M. Z. y Talley, W. K. (1980). Delphi stability and agreement criteria. *Technological forecasting and Social Change*, 16, 67-73.
- Daleva, M., Piperova-Dalbokova, D., Hadjiolova, I y Mincheva, L. (1982). Changes in the excretion of corticosteroids and catecholamines in tractor-drivers. *International Archives of Occupational Environmental Health*, 49(34), 345-352.
- Dalkey, N. C. (1969). *The Delphi method*, RM-5888. Santa-Monica, CA: RAND Corporation.
- Dalkey, N.C. (1972). The Delphi method: an experimental application of group opinion. En N. C. Dalkey, D. L. Rourke, R. Lewis y D. Snyder D (Eds.), *Studies in the quality of life* (pp. 13-54). Lexington Books, Lexington, MA, USA
- Dalkey, N. C.; Brown, B. y Cochran, S. (1970). The delphi method, III: Use of self rating to improve group estimates. *Technological forecasting and social change*, 1, 283-291.
- Dalkey, N.C. y Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, 9(3), 458-467.
- Dalkey, N. C. y Rourke, D. L. (1971). *Experimental assessment of Delphi procedures with group value judgments*. Santa Monica, CA: Rand Corporation.
- De Arquer, M. I., Daza, F. M. y Nogareda, C. (1995). *Ambigüedad y conflicto de rol. Notas Técnicas de Prevención edición electrónica (NTP- 388)*. España: Instituto Nacional de Seguridad e Higiene en el Trabajo (INSHT). Recuperado de http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/301a400/ntp_388.pdf
- De Beeck, R. y Hermans, V. (2000). *Research on work related low back disorders*. Brussels: Institute for Occupational Safety and Health.

- De Jong, M. G., Pieters, R. y Fox, J. P. (2010). Reducing social desirability bias through item randomized response. An application to measure underreported desires. *Journal of Marketing Research*, 47(1), 14–27. doi:10.1509/jmkr.47.1.14
- De Jonge, J., Bosma, H., Peater, R. y Siegrist, J. (2000). Job strain, effort-reward imbalance and employee well being: A large-scale cross-sectional study. *Social Science & Medicine*, 50, 1317-1327.
- De Keyser, V. y Hansez, I. (1996). Vers une perspective transactionnelle du stress au travail: Pistes d'évaluations méthodologiques. *Cahiers de médecine du travail*, 33(3), 133-144.
- De la Gándara, J., García, C., González-Corrales, R. y Álvarez, M.T. (2005). Estudio del burnout mediante el test CUBO (Cuestionario Urgente de Burn-Out). *Análisis de fiabilidad. Anales de Psiquiatría*, 25, 223-229.
- De la Guardia G., Gómez, M. L., Sandoval, E. y García, G. (2012). Sesgo o error de medición. En Villa A. *Epidemiología y Estadística en Salud Pública* (pp. 151-163). México. McGraw.Hill.
- DeLamater, J. y McKinney, K. (1982). Response-effects of question content. En W. Dijkstra y J. van der Zouwen (Eds.), *Response behaviour in the survey-interview*. New York: Academic Press.
- De Lange, A. H., Taris, T. W., Kompier, M.A., Houtman, I. L. y Bongers, P. M. (2003). The very best of the millennium: longitudinal research and the demand-control-(support) model. *J Occupational health psychology* 8(4), 282-305. doi:10.1037/1076-8998.8.4.282
- De Lange, A. H., Taris, T. W., Kompier, M. A., Houtman, I. L. y Bongers, P. M (2005). Different mechanisms to explain the reversed effects of mental health on work characteristics. *Scandinavian Journal of Work Environment and Health*, 31, 3–14.
- De las Cuevas, C. y González de Rivera. J. L. (1992). Autoinformes y Respuestas sesgadas. *Anales de Psiquiatría*, 8(9), 362-366
- Delbecq, A. L.; Van de Ven, A. y Gustafson, D. H. (1975). *Group Techniquesfor Program Planning: A Guide to Nominal Group and Delphi Processes*. Glenview, IL: Scoth-Foreman and Co.

- Delgado, C. (1998). El problema del sesgo en los test. Revisión histórica y cuestiones críticas. *Revista de Ciencias Sociales*, 80, 21–44.
- Delgado-Rodríguez, M. y Llorca, J. (2004). Bias. *Journal of Epidemiology and Community Health*, 58(8), 635–641. doi:10.1136/jech.2003.008466
- Del Hoyo, M. A. (2004). *Estrés laboral. Instituto Nacional de Seguridad e Higiene en el Trabajo*. Madrid: INSHT.
- Del Rio, I. Y. (2006). Estrés y sueño. *Revista Mexicana de Neurociencias*, 7(1), 15-20.
- Demerouti, E., Bakker, A. B., Nachreiner, F. y Schaufeli, W. B. (2001). The job demands resources model of burnout. *Journal of Applied Psychology*, 86, 499–512.
- Demerouti, E., Bakker, A. B., Vardakou, I. y Kantas, A. (2003). The convergent validity of two burnout instruments: A multitrait-multimethod analysis. *European Journal of Psychological Assessment*, 18, 296-307.
- De Meyrick, J. (2003). The delphi method and health research. *Health Education Journal*, 103, 7–16. doi:10.1108/09654280310459112
- De Vaus, D. A. (2002). *Surveys in Social Research*. 5th edition. Crows Nest, NSW: Allen and Unwin.
- DeVellis, R. F. (1991). *Scale Development: Theory and Applications*. London: Sage.
- Devereux, J. J. y Buckle, P. W. (julio, 2000). Adverse work stress reactions – A review of the potential influence on Work Related Musculoskeletal Disorders (WMSDs). *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44(30), 5--457-5-460. doi: 10.1177/154193120004403015
- Devereux, J. J., Vlachonikolis, I. G. y Buckle, P. W. (2002). Epidemiological study to investigate potential interaction between physical and psychosocial factors at work that may increase the risk of symptoms of musculoskeletal disorder of the neck and upper limb. *Occupational & Environmental Medicine*, 59(4), 269-277.

- Dicken, C. (1963). Good Impression, Social Desirability, and Acquiescence as Suppressor Variables. *Educational and Psychological Measurement*, 23(4), 699-720. doi: 10.1177/001316446302300406
- Díaz, D., Blanco, A., Horcajo, J. y Valle, C. (2007). La aplicación del modelo del estado completo de salud al estudio de la depresión. *Psicothema*, 19(12), 286–294.
- Di Martino, V., Hoel, H. y Cooper, C. L. (2003). *Preventing violence and harassment in the workplace*. Dublin: European Foundation for the Improvement of Living and Working Conditions.
- Djindjic, N., Jovanovic, J., Djindjic, B., Jovanovic, M. y Jovanovic, J. J. (2012). Associations between the occupational stress index and hypertension, type 2 diabetes mellitus, and lipid disorders in middle-aged men and women. *The Annals of Occupational Hygiene*, 56(9), 1051–1062. doi:10.1093/annhyg/mes059
- Dodge, B. J. y Clark, R. E. (1977). Research briefings: *Research on the Delphi technique*. *Educational Technology*, 17(4), 58-60.
- Dolan, S., García, S. y Diez Piñol, M. (2005). *Autoestima, estrés y trabajo*. Madrid: McGraw-Hill, Interamericana de España.
- Dollard, M. F., Skinner, N., Tuckey, M. R. y Bailey, T. (2007). National Surveillance of Psychosocial Risk Factors in the Workplace: An International Overview. *Work & Stress*, 21(1), 1-29.
- Donaldson, S. I. y Grant-Vallone, E. J. (2002). Understanding Self Report Bias in *Organizational Behavior Research*, 17(2), 245–261.
- Drossman, D. A. (1998). Presidential address: Gastrointestinal illness and the biopsychosocial model. *Psychosomatic Medicine*, 60(3), 258 – 267.
- D'Souza, R. M., Strazdins, L., Lim, L., Broom, D.H. y Rodgers, B. (2003). Work and health in a contemporary society: demands, control, and insecurity. *Journal of Epidemiology Community Health*, 57, 849-854.

- Duffy, L. C., Zielezny, M.A., Marshall, J. R., Byers, T. E., Weiser, M. M., Phillips, J. F.,... Graham, S. (1991). Relevance of major stress events as an indicator of disease activity prevalence in inflammatory bowel disease. *Behavioral Medicine*, 17(3), 101-110.
- Dupret, E., Bocerean, C., Teherani, M., Feltrin, M. y Pejtersen, J. H. (2012). Psychosocial risk assessment: French validation of the Copenhagen Psychosocial Questionnaire (COPSOQ). *Scandinavian journal of public health*, 40(5), 482-90.
- Dupuis, M., Meier, E., Capel, R. y Gendre, F. (2015). Measuring individuals' response quality in self-administered psychological tests: an introduction to Gendre's functional method. *Frontiers in Psychology*, 6, 629, 1-12. doi:10.3389/fpsyg.2015.00629
- Duro, A (2005). Calidad de vida laboral y psicología social de la salud laboral: hacia un modelo de componentes comunes para explicar el bienestar laboral psicológico y la salud mental laboral de origen psicosocial. Fundamentos teóricos. Revista del Ministerio de Trabajo y Asuntos Sociales. *Serie Economía y Sociología*, 56, 15-56.
- Easterbrook, J. A. (1959). The effects of emotion on cue utilisation and the organisation of behaviour. *Psychological Review*, 46, 183-207.
- Echeburúa, E., Amor, P. J. y Corral, P. (2003). Autoinformes y entrevistas en el ámbito de la psicología clínica forense: limitaciones y nuevas perspectivas. *Análisis y Modificación de Conducta*, 29, 503-522.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37(2), 90-93.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Edwards, A. L. (1964). Social desirability and performance on the MMPI. *Psychometrika*, 29, 295-308.
- Edwards, A. L. (1990). Construct validity and social desirability. *American Psychologist*, 45, 287-289.

- Einarsen, S., Hoel, H. y Notelaers, G. (2009). Measuring bullying and harassment at work: Validity, factor structure and psychometric properties of the Negative Acts Questionnaire – Revised. *Work and Stress*, 23(1), 24-44.
- Einarsen, S. y Raknes, B. I. (1997). Harassment at work and the victimization of men. *Violence and Victims*, 12, 247-263.
- Einarsen, S., Raknes, B. J. y Matthiesen, S. B. (1994). Bullying and harassment at work and their relationships to work environment quality. An exploratory study. *European Work and Organizational Psychologist*, 4, 381-401.
- Eller, N. H., Netterstrom, B., Gyntelberg, F., Kristensen, T. S., Nielsen, F., Steptoe, A. y Theorell, T. (2009). Work-related psychosocial factors and the development of ischemic heart disease: a systematic review. *Cardiology in Review*, 17, 83–97.
- Ellingson, J. E., Smith, D. B. y Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology*, 86, 122-133.
- Elliot, L. L. (1961). Effects of Item Construction and Respondent Aptitude on Response Acquiescence. *Educational and Psychological Measurement*, 21, 405-415.
- Elliot, T. R., Chartrand, J. M. y Harkins, S. W. (1994). Negative affectivity, emotional distress, and the cognitive appraisal of occupational stress. *Journal of Vocational Behaviour*, 45, 185-201.
- Embree, B. y Whitehead, P.C. (1991). Validity and reliability of self reported drinking behavior: dealing with the problem of response bias. *Journal of Studies on Alcohol*, 54, 334-344.
- Elovainio, M., Kivimäki, M., Steen, N. y Vahtera, J. (2004). Job decision latitude, organizational justice and health: Multilevel covariance structure analysis. *Social Science and Medicine*, 58, 1659–1669.
- Elovainio, M., Kivimäki, M., Vahtera, J., Virtanen, M. y Keltikangas-Jarvinen, L. (2003). Personality as amoderator in the relations between perceptions of organizational justice and sickness absence. *Journal of Vocational Behavior*, 63, 379-395.

- Engstrom, T., Hanse, J. J. y Kadefors, R. (1999). Musculoskeletal symptoms due to technical preconditions in long cycle time work in an automobile assembly plant: A study of prevalence and relation to psychosocial factors and physical exposure. *Applied Ergonomics*, 30(5), 443–453.
- Enríquez, F y Domínguez, A. (2010). Influencia de la Deseabilidad Social (DS) en Reportes de Capacitación. *Psicología Iberoamericana*, 18(1), 69-79. Recuperado de <http://www.redalyc.org/pdf/1339/133915936008.pdf>
- Epel, E. S., B. McEwen, T. Seeman, K. Matthews, G. Castellazzo, K. D. Brownell, J. Bell y J. R. Ickovics (2000). Stress and body shape: stress-induced cortisol secretion is consistently greater among women with central fat. *Psychosomatic medicine*, 62(5) ,623-632.
- Escobar-Pérez, J. y Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en Medición*, 6, 27-36.
- Espelt, A., Doval, E. y Viladrich, C. (2014). ¿Es necesario que los test sean equitativos? *Psico*, 45(4), 432–433.
- Espelt, A., Viladrich, C., Doval, E., Aliaga, J., García-Rueda, R. y Tárrega, S. (2014). Uso equitativo de tests en ciencias de la salud. *Gaceta Sanitaria*, 28(5), 408–410. doi:10.1016/j.gaceta.2014.05.001
- European Network for Workplace Health Promotion, ENWHP. (2011). *Declaración de Edimburgo sobre la promoción de la salud mental y el bienestar en el trabajo*. Recuperado de http://www.enwhp.org/fileadmin/downloads/Publications/Edinburgh_Declaration.pdf
- European Foundation for the Improvement of Living and Working Conditions, Eurofound. (2007). *Fourth European Working Conditions Survey (4EWCS)*. Office for Official Publications of the European Communities, Dublín. Recuperado de https://www.eurofound.europa.eu/sites/default/files/ef_files/pubdocs/2006/98/en/2/ef0698en.pdf

European Foundation for the Improvement of Living and Working Conditions, Eurofound. (2010). *Questionnaire for Fifth European Survey on Working Condition (5EWCT)*. Publications Office of the European Union, Luxembourg. Recuperado de https://www.eurofound.europa.eu/sites/default/files/ef_files/surveys/ewcs/2010/documents/masterquestionnaire.pdf

European Foundation for the Improvement of Living and Working Conditions, Eurofound. (2012). *Fifth European Working Conditions Survey (5EWCS)*, Publications Office of the European Union, Luxembourg. Recuperado de https://www.eurofound.europa.eu/sites/default/files/ef_files/pubdocs/2011/82/en/1/EF1182EN.pdf

European Foundation for the Improvement of Living and Working Conditions (Eurofound) y Agencia Europea para la Seguridad y la Salud en el Trabajo (EU-OSHA). (2014). *Psychosocial risks in Europe: Prevalence and strategies for prevention*. Luxemburgo: Oficina de Publicaciones de la Unión Europea. Recuperado de https://www.eurofound.europa.eu/sites/default/files/ef_publication/field_ef_document/ef1443en_0.pdf

European Foundation for the Improvement of Living and Working Conditions, Eurofound. (2015). *Violence and harassment in European workplaces: Causes, impacts and policies*. Dublin. Recuperado de https://www.eurofound.europa.eu/sites/default/files/ef_comparative_analytical_report/field_ef_documents/ef1473en.pdf

European Foundation for the Improvement of Living and Working Conditions, Eurofound. (2017). *Six European Working Conditions Survey (6EWCS)*. Publications Office of the European Union, Luxembourg. Recuperado de <http://www.oect.es/InshtWeb/Contenidos/Documentacion/FICHAS%20DE%20PUBLICACIONES/EN%20CATALOGO/GENERALIDAD/ENCT%202015.pdf>

European Heart Network (1998). *Expert Group On Psychosocial and Occupation Factors. Social Factors, Stress and Cardiovascular Disease Prevention in the European Union*. Brussels: European Heart Network.

- Eurostat (2017). *European Semester 2017: macroeconomic imbalances and government deficits*. Recuperado de https://ec.europa.eu/info/business-economy-euro/economic-and-fiscal-policy-coordination/european-economy-explained/graphs-economic-topics/european-semester-2017-macroeconomic-imbalances-and-government-deficits_en
- Evans, R.O. (1982). Clinical relevance of the Marlowe-Crowne Scale: A review and recommendations. *Journal of Personality Assessment*, 46, 415-425.
- Everson-Rose, S. A. y Lewis, T. T. (2005). Psychosocial factors and cardiovascular diseases. *Annual Review of Public Health*, 26, 469-500.
- Eysenck, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: University of London Press.
- Eysenck, H. J. (1988). Personality, stress and cancer: prediction and prophylaxis. *British Journal of Medical Psychology*, 61, 57-75.
- Eysenck, H. J. y Eysenck, S. B. G. (1964). *The Eysenck Personality Inventory manual*. London: Routledge.
- Eysenck, H. J. y Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. New York: Crane, Russak & Company.
- Farber, B. A. (1984). Stress and burnout in suburban teachers. *Journal of Educational Research*, 77, 325-331.
- Farmer, M. E., Locke, B. Z., Moscicki, E. K., Dannenberg, A. L., Larsen, D. B. y Radloff, L. S. (1988). Physical activity and depressive symptoms: The NHANES 1 epidemiologic follow-up study. *American Journal of Epidemiology*, 128, 1340-1351.
- Fernández de Castro, A y López P. A (2013). Validación mediante método Delphi de un sistema de indicadores para prever, diseñar y medir el impacto sobre el desarrollo local de los proyectos de investigación en el sector agropecuario. *Revista Ciencias Técnicas Agropecuarias*, 22(3), 54 - 60. Recuperado de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2071-00542013000300010

- Fernández, M., Aguayo V., Gómez L. E., Verdugo, M. A., Amor, A .M. y Castela, E., (2015) *Estudio delphi de una escala de calidad de vida en daño cerebral. Siglo Cero. 253 Anejo*. Castilla y León (España): Universidad de Salamanca.
- Fernández, R. (2010). *La productividad y el riesgo psicosocial o derivado de la organización del trabajo*. Alicante: Ed. Club Universitario.
- Fernández-Ballesteros, R. (2003). Self-report questionnaires. En S.N. Haynes & .M. Heiby (Eds.) *Comprehensive Handbook of Psychological Assessment*. Vol 3: Behavioral Assessment (pp. 194-221). New York: Wiley.
- Fernández-Ballesteros, R. (2013). Los autoinformes. En: R. Fernández- Ballesteros (ed.). *Evaluación Psicológica, Conceptos métodos y estudio de casos* (pp. 235-270). Madrid, Pirámide.
- Fernández-López, J. A. (2008). "Ganarás el pan con el sudor de tu frente...": Génesis, 3: 19. *Revista de administración sanitaria siglo XXI*, 6(3), 513-524.
- Fernández-López, J. A., Siegrist, J., Rödel, A. y Hernández-Mejía, R. (2003). El estrés laboral: un nuevo factor de riesgo. ¿Qué sabemos y qué podemos hacer? *Atención Primaria*, 31(8), 524b–526. doi:10.1157/13047737
- Fernández-López, J. A., Fernández-Fidalgo, E. y Siegrist, J. (2005). El trabajo y sus repercusiones en la salud. El modelo “Desequilibrio Esfuerzo-Recompensa-DER”. *Revista de Calidad Asistencial*, 20(3), 165-170.
- Fernández-Montalvo, J. y Echeburúa, E. (2006). Uso y abuso de los autoinformes en la evaluación de los trastornos de personalidad. *Revista de Psicopatología Y Psicología Clínica*, 11(1), 1–12.
- Fernández- Seara, J. L. (1992). *Escala de Apercepción del Estrés*. Madrid: TEA Ediciones.
- Ferradans, C. (2008). Prevención de riesgos psicosociales. En *Análisis de la regulación jurídico-laboral de los riesgos psicosociales* (pp. 107-129). Recuperado de <http://www.seg-social.es/prdi00/groups/public/documents/binario/100617.pdf>

- Ferrán, M. (2001). *SPSS para Windows. Análisis estadístico*. Madrid: Osborne McGraw-Hill.
- Ferrando, P. J. (2005). Factor analytic procedures for assessing social desirability in binary items. *Multivariate Behavioral Research*, 40(3), 331-349.
- Ferrando, P. J. y Anguiano-Carrasco, C. (2009). Assessing the impact of faking on binary personality measures: An IRT-Based Multiple-Group Factor Analytic Procedure. *Multivariate Behavioral Research*, 44(4), 497-524.
- Ferrando, P. J. y Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- Ferrando, P. J. y Anguiano-Carrasco, C. (2012). Response Certainly, Conscientiousness, and Self-concept Clarity as antecedents of Acquiescence: A prediction model. *Anuario de Psicología*, 42(1), 103 - 112. Recuperado de <http://www.redalyc.org/pdf/970/97024524007.pdf>
- Ferrando, P. J. y Chico, E. (2000). Adaptación y análisis psicométrico de la escala de deseabilidad social de Marlowe y Crowne. *Psicothema*, 12, 383-389.
- Ferrando, P. J., Condon, L. y Chico, E. (2004). The convergent validity of acquiescence: an empirical study relating balanced scales and separate acquiescence scales. *Personality and individual differences*, 37(7), 1331-1340.
- Ferrando, P. J., Demestre, J., Anguiano-Carrasco, C. y Chico, E. (2011). An IRT assessment of Rotter I-E scale: A new approach and some considerations. *Psicothema*, 23(2), 282-288.
- Ferrando, P. J. y Lorenzo-Seva, U. (2010). Acquiescence as a source of bias and model and person misfit: a theoretical and empirical analysis. *The British Journal of Mathematical and Statistical Psychology*, 63(2), 427-448. doi: 10.1348/000711009X470740
- Ferrando, P. J., Lorenzo-Seva, U. y Chico, E. (2009). A general factor-analytic procedure for assessing response bias in questionnaire measures. *Structural Equation Modeling*, 16, 364-381.
- Ferrer, R. (2010). La evaluación de riesgos psicosociales en tiempos de crisis. *Revista de Prevención de Riesgos Psicosociales y bienestar en el trabajo*, 2, 67-85.

- Ferrer, R., Guilera, G. y Peró, M. (2011). *Propiedades Psicométricas del Instrumento de Valoración de Riesgos Psicosociales del Instituto Nacional de Seguridad e Higiene en el Trabajo (FPSICO)*. Barcelona: Universitat de Barcelona.
- Ferrie, J. E. (1999). Health consequences of job insecurity. En J.E. Ferrie, M.G.Marmot y J.Griffiths (Eds.), *Labour market changes and job insecurity: a challenge for social welfare and health promotion* (pp. 59-100). Copenhagen, Dinamarca: World Health Organization.
- Ferrie, J. E., Shipley, M. J., Stansfeld, S. A. y Marmot, M. G. (2002). Effects of chronic job insecurity and change in job security on self-reported health, minor psychiatric morbidity, physiological measures, and health related behaviours in British civil servants: the Whitehall II study. *Journal of Epidemiology and Community Health*, 56, 450-454.
- Feuerstein, M., Shaw, W. S., Nicholas, R. A. y Huang, G. D. (2004). From confounders to suspected risk factors: psychosocial factors and work-related upper extremity disorders. *Journal of electromyography and kinesiology*, 14, 171-8.
- Fidalgo, Á. M., Mellenbergh, G. J. y Muñoz, J. (1998). Comparación del procedimiento Mantel-Haenszel frente a los modelos loglineales en la detección del funcionamiento diferencial de los ítems. *Psicothema*, 10(1), 209–218.
- Fidalgo, Á. M. y Paz, M. D. (1995). Modelos lineales logarítmicos y funcionamiento diferencial de 10s ítems. *Anuario de psicología*, 64, 57-66.
- Fidalgo, M. (2006a). *Síndrome de estar quemado por el trabajo o “burnout” (I): definición y proceso de generación*. NTP 704. Madrid. (INHT).
- Fidalgo, M. (2006b). *Síndrome de estar quemado por el trabajo o “burnout” (II): consecuencias, evaluación y prevención*. NTP 705. Madrid. (INHT).
- Fidalgo, A. M. y Piñuel, I. (2004). La escala Cisneros como herramienta de valoración del mobbing. *Psicothema*, 16, 615-624.
- Fielden, S. L. y Peckar, C. J. (1999). Work stress and hospital doctors: a comparative study. *Stress Medicine*, 15(3), 137-141.

- Finlay, W. M. L. y Lyons, E. (2002). Acquiescence in interviews with people with mental retardation. *Mental Retardation*, 40(1), 14-29. doi:10.1352/0047-6765(2002)040<0014:AIWPW>2.0.CO;2
- Fischer, A. (2000). *Gender and Emotions: Social psychological perspectives*. Cambridge University Press.
- Fleming, P. y Zizzo, D. J. (2011). Social desirability, approval and public good contribution. *Personality and Individual Differences*, 51, 258-262. doi: 10.1016/j.paid.2010.05.028
- Folkman, S. y Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior*, 21, 219-239.
- Forgas, J. P., Chan, N. Y. M. y Laham, S. M. (2001). Affective influences on thinking and behavior: Implications for clinical, applied and preventive psychology. *Applied and Preventive Psychology: Current Scientific Perspectives*, 10(4), 225-242
- Forgas, J. R. y George, J. M. (2001). Affective influences on judgments and behavior in organizations: An information processing perspective. *Organizational Behavior and Human Decision Processes*, 86, 3-34.
- Fornés, J., Martínez-Abascal, M. A. y García de la Banda, G. (2008). Análisis factorial del Cuestionario de Hostigamiento Psicológico en el Trabajo en profesionales de Enfermería. *Internacional Journal of Clinical and Health Psychology*, 8, 267-383.
- Fowler-Brown, A. G., Bennett, G. G., Goodman, M. S., Wee, C. C., Corbie-Smith, G. M. y James, S. A. (2009). Psychosocial Stress and 13-year BMI Change Among Blacks: The Pitt County Study. *Obesity (Silver Spring)* 17(11), 2106–2109. doi: 10.1038/oby.2009.130
- Freeman, L. M. Y. y Gil, K. M. (2004). Daily stress, coping, and dietary restraint in binge eating. *International Journal of Eating Disorders*, 36, 204-212.
- French, J. R. P., Jr., Caplan, R. D. y Van Harrison, R. (1982). *The mechanisms of job stress and strain*. Chichester: Wiley.
- French, B. F. y Maller, S. J. (2007). Iterative purification and effect size use with logistic regression for differential item functioning detection. *Educational and Psychological Measurement*, 67(3), 373-393

- Frese, M. (1985). Stress at work and psychosomatic complaints: A causal interpretation. *Journal of Applied Psychology*, 70, 314-328.
- Frese, M. y Zapf, D. (1994). Action as the core of work psychology: A German approach. En H.C. Triandis, M.D. Dunnette & J.M. Hough (Eds.), *Handbook of industrial and organizational psychology*. Palo alto, CA: Consulting Psychologists.
- Freudenberger H. J. (1974). Staff Burn-Out. *Journal of Social Issues*, 30(1), 159-166.
- Freudenberger, H. J. y Richelson, G. (1980). *Burnout: The high cost of high achievement*. New York: Anchor Press.
- Fröberg, J., Karlsson, C. G., Levi, L., Lidberg, L. y Seeman, K. (1970). Conditions of work: psychological and endocrine stress reactions. *Archives of Environmental Health*, 21, 789-797. doi: 10.1080/00039896.1970.10667333
- Frutos, B. D. (2013). *La construcción de indicadores válidos en la investigación en comunicación*. Universidad de Valladolid: Facultad de Ciencias Sociales, Jurídicas y de la Comunicación (UVa] Segovia).
- Fukudo, S. y Suzuki, J. (1987). Colonic motility, autonomic function, and gastrointestinal hormones under psychological stress on irritable bowel syndrome. *The Tohoku Journal of Experimental Medicine*, 151, 373-385.
- Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences*, 7, 385-400.
- Gamero-Burón, C. y González, M. L. (2013). *Costes Socio-Económicos de los Riesgos Psicosociales. Balance de situación actual y propuestas metodológicas para avanzar en su medición o cálculo*. Observatorio permanente de riesgos psicosociales UGT - Universidad de Málaga - INSHT.
- Ganster, D. C., Fox, M. L. y Dwyer, D. J. (2001). Explaining employees' health care costs: a prospective examination of stressful job demands, personal control, and physiological reactivity. *The Journal of applied psychology*, 86(5), 954-964. doi: 10.1037/0021-9010.86.5.954

- Ganster, D. C. y Rosen, C. R. (2013). Work stress and employee health: A multidisciplinary review. *Journal of Management*, 39(5), 1085-112. doi: 10.1177/0149206313475815
- Ganster, D. C. y Schaubroeck, J. (1991). Work, stress and employee health. *Journal of Management*, 17, 235-271.
- García, E. M. y Magaz, A. (2009). *¿Cómo valorar tests psicométricos? Errores conceptuales y metodológicos en la evaluación psicoeducativa*. Vizcaya: Grupo Albor-Cohs.
- García, L. y Fernández, S. (2008). Procedimiento de aplicación del trabajo creativo en grupo de expertos. *Energética*, XXIX (2), 46-50.
- García, M. y Olarte, S (2010). La regulación de los riesgos psicosociales en España: experiencia y prospectiva". *Revista de Prevención de Riesgos Psicosociales y Bienestar en el Trabajo* 1, 13-43.
- García-Izquierdo, M. (1991a). Burnout en profesionales de enfermería de centros hospitalarios. *Revista de Psicología del Trabajo y de las Organizaciones*, 7(8), 3-12.
- García-Izquierdo, M. (1991 b). Estrés en contextos laborales. *Comunicación presentada a las III Jornadas de Psicología clínica y salud*. Murcia.
- García-Izquierdo, M., Llor, B., Sáez, M. C., Ruiz, J. A., Blanco, J. R. y Campillo, M. J. (2004). Evaluación del acoso psicológico en el trabajo: El NAQ-RE. Revisión de la adaptación española. *Comunicación oral presentada al VII European Conference on Psychological Assessment*. Málaga. España.
- García-Layunta, M., Tomás, J. M. y Oliver, A. (2005). Riesgos psicosociales y bienestar en el trabajo la satisfacción laboral desde el modelo Warr. *Revista de la Sociedad Española de Medicina y Seguridad del Trabajo*, 1(1),87-94.
- Garrido, A. y Álvaro, J.L. (2007). *Psicología social. Perspectivas psicológicas y sociológicas*. (2ª Edición). Madrid: McGraw-Hill.
- Garrido-Pinzón, J., Uribe-Rodríguez, A. F. y Blanch, J. M. (2011). Riesgos psicosociales desde la perspectiva de la calidad de vida laboral. *Acta colombiana de Psicología*, 14(2), 27-34.

- Gelin, M. A. y Zumbo, B. D. (2003). Differential item functioning results may change depending on how an item is scored: An illustration with the centre for epidemiologic studies depression scale. *Educational and Psychological Measurement*, 63(1), 65-74.
- George, D. y Mallery, P. (2007). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (7th ed.). Boston: Allyn y Bacon.
- Gianoulakis, C. (1998). Alcohol-Seeking Behavior: The roles of the hypothalamic-pituitary-adrenal axis and the endogenous opioid system. *Alcohol Health & Research World*, 22(3), 202-210.
- Gibson, E. L. (2006). Emotional influences on food choice: sensory, physiological and psychological pathways. *Physiology & Behavior*, 89, 53-61.
- Gil-Gómez, B. y Pascual-Ezama D. (2012). La metodología Delphi como técnica de estudio de la validez de contenido. *Anales de Psicología*, 28(3), 1011-1020.
- Gillespie, D. (1980). Correlates for active and pasives Burnout, types of Burnout. *Journal of social service Research*, 4(2), 1-16.
- Gil-Monte, P. R. (2003). Redalyc.Burnout syndrome: ¿síndrome de quemarse por el trabajo, desgaste profesional, estrés laboral o enfermedad de Tomás? *Revista de Psicología Del Trabajo Y de Las Organizaciones*, 19(2), 181-197.
- Gil-Monte, P. R. (2005). *El síndrome de quemarse por el trabajo (Burnout): una enfermedad laboral en la sociedad del bienestar*. Madrid, España: Pirámide.
- Gil-Monte, P. R. (2009). Algunas razones para considerar los riesgos psicosociales en el trabajo y sus consecuencias en la salud pública. *Revista Española de Salud Pública*, 83(2), 169-173.
- Gil-Monte, P. R. (2010). Situación actual y perspectiva de futuro en el estudio del estrés laboral: la Psicología de la Salud Ocupacional. *Información psicológica*, 100, 68-83.
- Gil-Monte, P. R. (2014). Psicosociología aplicada al estudio de los riesgos laborales: estrés y estrategias de afrontamiento. En Gil-Monte, P.R., coord. *Manual de Psicosociología aplicada al trabajo y a la prevención de riesgos laborales* (pp. 25-48.) Madrid: Pirámide.

- Gil-Monte, P. y Peiró, J. M. (1997). *Desgaste psíquico en el trabajo*. Madrid: Síntesis Psicología.
- Gimeno, D. (2004). Los factores de riesgo laboral de naturaleza psicosocial y su prevención. *Archivos de Prevención de Riesgos Laborales*, 7(3), 119–120.
- Giner, C. A. (2012). Evaluación de riesgos psicosociales en el trabajo. *Anales Del Derecho*, 30, 254-296. doi:10.6018/analesderecho
- Glass, J. y Fujimoto, T (1994). Housework, paid work, and depression among husbands and wives. *Journal of Health and Social Behavior*, 35(2) 179-191.
- Glasscock, D. J., Rasmussen, K., Carstensen, O. y Hansen, O. N. (2006). Psychosocial factors and safety behaviour as predictors of accidental work injuries in farming. *Work & Stress*, 20(2), 173-189.
- Gluck, M. E. (2006). Stress response and binge eating disorder. *Appetite*, 46(1), 26–30.
- Godoy-Izquierdo, D. D., Martínez, A. A. y Godoy, J. F. (2008). La "Escala de Balance Afectivo". Propiedades psicométricas de un instrumento para la medida del afecto positivo y negativo en población española (The "Affect Balance Scale": Its psychometric properties as a tool for measuring positive and negative affect in the Spanish population). *Clínica y Salud*, 19, 157-189.
- Goffin, R. D. y Boyd, A. C. (2009). Faking and personality assessment in personnel selection: Advancing models of faking. *Canadian Psychology*, 50(3), 151-160.
- Gómez-Benito, J. e Hidalgo, M. D. (1997). Evaluación del funcionamiento diferencial en ítems dicotómicos: una revisión metodológica. *Anuario de Psicología*, (74), 3–32.
- Gómez-Benito, J., Hidalgo, M. D. y Ferré, G. G. (2005). Introducción al funcionamiento diferencial del ítem. *Revista Psicología.com*, 9(2). Recuperado de <http://psiqui.com/2-9864>
- Gómez-Benito, J., Hidalgo, M. D. y Guilera, G. (2010). El sesgo de los instrumentos de medición. Tests justos. *Papeles del Psicólogo*, 31(1), 75-84.
- Gómez-Biedma, S., Vivó, M. y Soria, E. (2001). Pruebas de significación en Bioestadística. *Revista de Diagnóstico Biológico*, 50(4), 207-218.

- González, A., Guerrero, A. y Maldonado, J. A. (2001) Desarrollo de un sistema experto para la aplicación del método Delfos (delphi). *BEIO, Boletín de Estadística e Investigación Operativa*, 17(4), 9-10.
- González, H. e Iruarrizaga, I. (2005). Evaluación de las distorsiones de respuesta mediante el MMPI-2. *Papeles del Psicólogos*, 26, 129-137.
- González de Rivera, J. L (1990). *El Índice de Reactividad al Estrés*. Madrid: Inteva.
- González de Rivera, J. L. y Rodríguez-Abuín, M. (2003). Cuestionario de estrategias de acoso psicológico: El LIPT-60 (Leymann Inventory of Psychological Terrorization) en versión española. *Psiquis*, 24, 59-66.
- González-Gutiérrez, J. L., Peñacoba, C., Velasco, L., López-López, A., Mercado, F. y Barjola, P. (2009). Recursos cognitivos de percepción de control, procesos de estrés-recuperación y malestar afectivo en la fibromialgia. *Psicothema*, 21(3), 359-368.
- González-Llaneza F. M. (2007). *Instrumentos de Evaluación psicológica*. La Habana: Editorial Ciencias Médicas.
- González-Romá, V., Väänänen, A., Ripoll, P., Caballer, A., Peiró, J.M. y Kivimäki, M. (2005). *Psychological climate, sickness absence and gender*. *Psicothema*, 17(1), 169-174.
- González-Trijueque, D. y Delgado, M. (2008). El acoso psicológico en el lugar de trabajo, antecedentes organizacionales *Boletín de Psicología*, 93(9), 7-20.
- González-Trijueque, D., Giachero, S. y Delgado, S. (2012). Riesgos psicosociales en el lugar de trabajo: aproximación teórica y marco legal en Uruguay. *Ciencias Psicológicas*, VI(1), 75–87.
- Goodman, L. A. (1961), Snowball sampling: *The Annals of Mathematical Statistics*, 32(1), 148-170.
- Gordon, R. A. (1987). Social Desirability Bias: A demonstration and technique for its reduction. *Teaching of Psychology*, 14(1), 40-42.
- Gordon, T. J. (1994). *The Delphi method*. Washington, DC: American Council for the United Nations University.

- Gordon, T. J. (2003). The Delphi Method. En J. C. Glenn & T. J. Gordon (Eds.), *Futures research methodology V2.0*. Washington D.C.: American Council for the United Nations University.
- Gordon, T. J. (2009a). *The Delphi Method*. En J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology - V.3.0*. The Millennium Project, American Council for the United Nations University
- Gordon, T. J. (2009b). The real - time Delphi method. En J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology - V.3.0*. The Millennium Project, American Council for the United Nations University
- Gordon, T. J. y Helmer, O. (1964) *Report on a Long Range Forecasting Study*. R-2982. Santa Mónica: The Rand Corporation,
- Gorina, A. y Alonso, I. (2013.). Un sistema de procedimientos metodológicos para perfeccionar el procesamiento de la información en las investigaciones sociales. *Revista Didasc@lia: D&E*, 3(6), 91-108.
- Gosling, S. D., Vazire, S., Srivastava, S. y John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, 59(2), 93- 104. Recuperado de <http://darkwing.uoregon.edu/~sanjay/pubs/webstudies.pdf>
- Gough, H. G. (1952). On making good impression. *Journal of Educational Research*, 46, 33-42.
- Gough, H. G. y Heilbrun, A. B. (1983). *Adjective Check List manual* (rev. ed.). Palo Alto, CA: Consulting Psychologists Press.
- Goulet, J. P., Lund, J. P., Montplaisir, J. Y. y Lavigne, G. J. (1993). Daily clenching, nocturnal bruxism, and stress and their association with TMD symptoms. *Journal of Orofacial Pain*, 7, 120-127.
- Gove, W. R. y Geerken, M. R. (1977). Response bias in surveys of mental health: An empirical investigation. *American Journal of Sociology*, 82, 1289–1317.

- Gracia-Camón, D. A. (s.f.). *Preguntas frecuentes de la metodología de evaluación de Riesgos Psicosociales INERMAP*. Instituto MAPFRE. Recuperado de <http://www.inermap.com/software/imagenes/FAQ.pdf>
- Gracia-Camón, D. A. (2003). Criterios objetivos para la evaluación de la carga de trabajo derivada de factores psicosociales y organizacionales. *Mapfre Seguridad*, 23(90), 29-37.
- Gracia-Camón, D. A. (2006). Método del Instituto de Ergonomía MAPFRE (INERMAP). En Meliá, J.L., Nogareda, C., Lahera, M., Duro, A., Peiró, J.M., Pou, R. et al. (Eds.), *Perspectivas de Intervención en Riesgos Psicosociales. Evaluación de Riesgos* (pp. 86-103). Barcelona: Foment del Treball Nacional.
- Grau, A., Flichtentrei, D., Suñer, R., Prats, M. y Braga, F. (2009). Influencia de factores personales, profesionales y transnacionales en el síndrome de burnout en personal sanitario hispanoamericano y español (2007). *Revista Española de Salud Pública*, 83(2), 215-230.
- Green, B., Jones, M., Hughes, D. y Williams, A. (1999). Applying the Delphi Technique in a Study of GP's Information Requirements. *Health and Social Care in the Community*, 7(3), 198-205.
- Green, K. L. y Johnson, J. V. (1990). The effect of psychological work organization on patterns of cigarette smoking among male chemical plant employees. *American Journal of Public Health*, 80(11), 1368-1371.
- Greenleaf, E. A. (1992). Measuring extreme response style. *Public Opinion Quarterly*, 56, 328–351.
- Greeno, C. G. y Wing, R. R. (1994). Stress-induced eating. *Psychological Bulletin*, 115, 444–464.
- Gregory, R. (2012). *Pruebas Psicológicas: historia, principios y aplicaciones*. México. Ed. Pearson.
- Grimm, S. D. y Church, A.T. (1999). A cross-cultural study of response biases in personality measures. *Journal of Research in Personality*, 33(4), 415-441.

- Guàrdia, J. (2008). *Evaluación de los principales métodos de evaluación de riesgos psicosociales*. Facultad de Psicología. Universidad de Barcelona. Recuperado de https://www.juntadeandalucia.es/empleo/webiapr/larpsico/sites/es.empleo.webiapr.larpsico/files/recursos/documentacion_normativa/VALORACIONMETODOSDEEVALUACIONRIESGOSpdf.pdf
- Guàrdia, J. (Coord.) (2010a). *La evaluación de riesgos psicosociales. Guía de buenas prácticas. Métodos de evaluación y sistemas de gestión de riesgos psicosociales: un balance de utilidades y limitaciones*. Jaén: Blanca Impresores, S.L.
- Guàrdia, J. (2010b). Taller de Prevención de Riesgos Psicosociales. *Revista de Prevención de Riesgos Psicosociales y Bienestar en el Trabajo*, 1, 171-192. Recuperado de https://www.juntadeandalucia.es/empleo/webiapr/larpsico/sites/es.empleo.webiapr.larpsico/files/recursos/revista_cientifica/RevistaLarpsico01.pdf
- Guàrdia, J. y Però, M. (2010). La evaluación de riesgos de origen psicosocial, ¿Una cuestión de método, de técnica, de instrumentos, de improvisación? En M. Rufino, C. Molina y E. González (Eds.). *Anuario internacional sobre prevención de riesgos psicosociales y calidad de vida en el trabajo* (pp. 89-110). Jaén: Blanca Impresores, S.L. Recuperado de <http://portal.ugt.org/saludlaboral/observatorio/publicaciones/anuarios/anuario2010.pdf>
- Guàrdia, J., Però, M. y Barrios, M. (2008). Propiedades psicométricas de la batería de evaluación de riesgos psicosociales en la mediana y pequeña empresa. *Psicothema*, 20(4), 939–944.
- Guélaud, F, Beauchesne, M.N., Gautrat, J. y Roustang, G. (1978). *Para un análisis de las condiciones del trabajo obrero en la empresa. Método LEST*. Centro Nacional de Investigación. París: Laboratoire d'Économie et de Sociologie de Travail. CNRS
- Guest, D. (2004). The Psychology of the Employment Relationship: An Analysis Based on the Psychological Contract. *Applied Psychology*, 53(4), 541-555.
- Guest, D. y Conway, N. (2002). Communicating the psychological contract: an employer perspective. *Human Resource Management Journal*, 12(2), 22-38.

- Guillén-Riquelmeé, A. y Buela-Casal, G. (2011). Actualización psicométrica y funcionamiento diferencial de los items en el State Trait Anxiety Inventory (STAI). *Psicothema*, 23(3), 510–515.
- Gupta, U. y Clarke, R. (1996). Theory and Applications of the Delphi Technique: A Bibliography (1975-1994). *Technological Forecasting and Social Change*, 53(2), 185-211
- Gutiérrez, S., Sanz, J., Espinosa, R., Gesteira, C. y García-Vera, M. P. (2016). La Escala de Deseabilidad Social de Marlowe-Crowne: baremos para la población general española y desarrollo de una versión breve. *Anales de Psicología*, 32, 206-217.
- Hacker, W., Skell, W. y Straub, W. (1968). *Arbeitspsychologie und wissenschaftlich- technische revolution*. Berlin: Deutscher Verlag der Wissenschaften.
- Hackman, J. R. y Oldham, G.R. (1975). Development of the Job Diagnostic Survey. *Journal of Applied Psychology*, 60, 159-170.
- Hackman J. R. y Oldham, G.R. (1980). *Work redesigning*. Reading Massachusetts: Adisson-Wesley.
- Hagen, K. B., Magnus, P. y Vetlesen, K. (1998). Neck/shoulder and low-back disorders in the forestry industry: Relationship to work tasks and perceived psychosocial job stress. *Ergonomics*, 41(10), 1510–1518.
- Hägglund, G. (1982). Factor analysis by instrumental variable methods. *Psychometrika*, 47, 209-222
- Hair, J., Anderson, R., Tatham, R. y Black, W. (2007). *Análisis Multivariante*. 5ta Ed. Prentice Hall Iberia, Madrid. España.
- Hamberger, L. K. y Lohr, J. M. (1984). *Stress and Stress Management: Research and applications* New York: Springer publishing company.
- Hamilton, D. L. (1968). Personality attributes associated with extreme response style. *Psychological Bulletin*, 69(3), 192-203.
- Hansson, M., Boström, C. y Harms-Ringdahl, K. (2006). Sickness absence and sick-ness attendance – What people with neck or pain think? *Social Science and Medicine*, 62, 2183-2195.

- Harding, S. D. (1982). Psychological well-being in Great Britain: An evaluation of the Bradburn Affect Balance Scale. *Personality and Individual Differences*, 3, 167-175.
- Harrma, M. (2006). Work hours in relation to work stress, recovery and health. *Scandinavian Journal of Work Environment and Health*, 32(6), 502-514.
- Hart, S. G. y Wickens, C. D. (1990). Workload assessment and prediction. En H. R. Booher (Ed.), *MANPRINT: An emerging technology. Advanced concepts for integrating people, machines and organizations* (pp. 257-300). Nueva York: Van Nostrand Reinhold.
- Hartshorne, H. y May, M. A. (1928). *Studies in the Nature of Character* (Vol. I. Studies in Deceit.). New York, NY: MacMillan.
- Harvey, R., Billings, R. y Nilan, K. (1985). Confirmatory factor analysis of the Job Diagnostic Survey: Good news and bad news. *Journal of Applied Psychology*, 70, 461-468.
- Harzing, A.W., Brown, M., Köster, K. y Zhao, S. (2012). Response style differences in cross-national research: dispositional and situational determinants, *Management International Review*, 52(3), 341- 363. Recuperado de https://harzing.com/download/rsdi_spo.pdf
- Hasson, F. y Keeney, S. (2011) Enhancing rigour in the Delphi technique research. *Technological Forecasting and Social Change*, 78, 1695-1704.
- Hathaway, S. y McKinley, F. (1943). *The Minnesota Multiphasic Personality Inventory*. Nueva York: Psychological Corporation.
- Hathaway, S. R. y McKinley, J. C. (1951). *The MMPI manual*. New York: The Psychological Corporation.
- Haynes, C. E., Wall, T. D., Bolden, R. I. y Rick, J. E. (1999). Measures of perceived work characteristics for health services research: Test of a measurement model and normative data. *British Journal of Health Psychology*, 4, 257-275.
- Head, J., Martikainen, P., Kumari, M., Kuper, H. y Marmot, M. (2002). *Work Environment, Alcohol Consumption and Ill-Health: The Whitehall II Study*. Suffolk, UK: HSE Books. Recuperado de http://www.hse.gov.uk/research/crr_pdf/2002/crr02422.pdf

- Health and Safety Professionals Alliance (HaSPA). (2012). *The Core Body of Knowledge for Generalist OHS Professionals*. Tullamarine, VIC. Safety Institute of Australia.
- Heaven, P. C. L. (1983). Authoritarianism or acquiescence? South African findings. *Journal of Social Psychology*, 119, 11–15.
- Heinemann. P. (1972). *Mobbing-Gruppvoald bland barn och vuxna IMobbing-group violence by children and adults*. Stockholm: Natur och Kultur.
- Heinisch, D. C. y Jex, S. M. (1997). Negative affectivity y gender as moderators of the relationship between work-related stressors y depressed mood at work. *Work & Stress*, 11, 46-57.
- Heinrich H. W. (1931). *Industrial accident prevention: a scientific approach*. New York: McGraw-Hill.
- Hellerstedt, W. L. y Jeffrey, R. W. (1997). The association of job strain and health behaviors in men and women. *International Journal of Epidemiology*, 26, 575–583.
- Herche, J. y Engelland, B. (1996). Reversed-Polarity Items and Scale Dimensionality. *Journal of the Academy of Marketing Science*, 24(4), 366-374. doi: 10.1177/0092070396244007
- Hernández, M., Garrido, F. y Salazar, E. (2002). Sesgos en estudios epidemiológicos. *Salud Pública de México*, 42(5), 438-446.
- Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación*. 6ªed. México: Editorial Mcgraw–Hill.
- Hidalgo, M. D. y Gómez, J. (2000). Comparación de la eficacia de regresión logística politómica y análisis discriminante logístico en la detección del DIF no uniforme. *Psicothema*, 12(SUPPL. 2), 298–300.
- Hidalgo, M. D. y Gómez, J. (2006). Nonuniform DIF detection using discriminant logistic analysis and multinomial logistic regression: A com parison for polytomous items. *Quality & Quantity*, 40, 805-823
- Hidalgo, M. D., Gómez, J. y Padilla, J. L. (2005). Regresión logística: alternativas de análisis en la detección del funcionamiento diferencial del ítem. *Psicothema*, 17(3), 509-515.

- Hidalgo, M.D. y López-Pina, J.A. (2004). Differential item functioning detection and effect size: A comparison between logistic regression and Mantel-Haenszel procedures. *Educational and Psychological Measurement*, 64, 903-915. doi: 10.1177/0013164403261769
- Higgins, J. P. T. y Green, S. (eds). (2011). *Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0*. The Cochrane Collaboration. Recuperado de <http://handbook-5-1.cochrane.org/>
- Hirigoyen, M. (2013a). *El acoso moral en el trabajo: distinguir lo verdadero de lo falso*. Buenos Aires: Paidós.
- Hirigoyen, M. (2013b). *El acoso moral: el maltrato psicológico en la vida cotidiana*. Barcelona: Paidós Ibérica, S.A.
- Hoel, H., Cooper, C.L. y Faragher, B. (2001). The experience of bullying at work in Great Britain: The impact of organisational status. *European Journal of Work and Organisational Psychology*, 10, 414-425
- Hogan, H. W. (1972). Fakability of the Adorno F-scale. *Psychological Reports*, 30, 15-21.
- Hogan, R. y Nicholson, R.A. (1988). The meaning of personality test scores. *American Psychologist*, 43, 621-626.
- Holden, R. R. (2007). Socially desirable responding does moderate personality scale validity both in experimental and in nonexperimental contexts. *Canadian Journal of Behavioral Science*, 39(3), 184-201.
- Holden, R. R. (2008). Underestimating the effects of faking on the validity of self-report personality scales. *Personality & Individual Differences*, 44(1), 311-321.
- Holland, W. P. y Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. En H. Wainer y H. I. Braun (Eds.), *Test validity* (pp. 129- 145). Hillsdale NJ: Erlbaum.
- Holmes, T. y Rahe. R. (1967). Escala de reajuste social. *Revista de Investigación Psicosomática*, 11, 213-218.

- Hoogendoorn, W. E., van Poppel, M. N. M., Bongers, P. M., Koes, B. W. y Bouter, L. M. (2000). Systematic review of psychosocial factors at work and private life as risk factors for back pain, *SPINE*, 25(16), 2114-2125.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Horowitz, M. J. (1976). *Stress response syndromes*. Nueva York: Jason Aronson.
- Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D. y McCloy, R. A. (1990). Criterion-Related Validities of Personality Constructs and the Effect of Response Distortion on Those Validities. *Journal of Applied Psychology*, 75(5), 582-595. doi:10.1037/0021-9010.75.5.581
- Howard, G. S. (1994). Why do people say nasty things about self-reports? *Journal of Organizational Behavior*, 15, 399-404. doi: 10.1002/job.4030150505
- Hsu, C. C. y Sandford, B. A. (2007). The Delphi Technique: Making Sense of Consensus. *Practical Assessment Research & Evaluation*, 12(10). Recuperado de <http://pareonline.net/getvn.asp?v=12&n=10>
- Hu, L. y Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Huh, J., DeLorme, D. E. y Reid, L. N. (2006). Perceived third-person effects and consumer attitudes on prevetting and banning DTC advertising. *Journal of Consumer Affairs*, 40(1), 90-116.
- Hui, C. H. y Triandis, H. C. (1985). The instability of response sets. *Public Opinion Quarterly*, 49, 253-260.
- Hui. C. H. y Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, 20(3), 296-309.
- Hurrell, J. J. Jr. y McLaney, M. A. (1988). Exposure to job stress- A new psychometric instrument. *Scandinavian Journal of Work Environment and Health*, 14, 27-27.

- Idzikowski, C. y Baddeley, A. D. (1983). Fear and Dangerous Environments, in R. Hockey (Ed.), *Stress and Fatigue in Human Performance* (pp. 123-144). Chichester, N.Y.: Willey.
- INSHT (1991). *Condiciones de trabajo: instrucciones de uso*. Traducción de “Conditions de travail, mode d’emploi” Agence Nationale pour l’Amélioration de Conditions de Travail (ANACT). Madrid: INSHT.
- INSHT, IBV (2003). *Manual para la evaluación y prevención de riesgos ergonómicos y psicosociales en PYME*. [Coordinadores: García-Molina, C., Villar, M.F.] Instituto de Biomecánica de Valencia (IBV), Instituto Nacional de Seguridad e Higiene en el Trabajo (INSHT), Ministerio de Trabajo y Asuntos Sociales, Madrid.
- Instituto Nacional de Estadística [INE]. (2016). *Encuesta de población activa (EPA), tercer trimestre*. Nota de prensa. Recuperado de <http://www.ine.es/daco/daco42/daco4211/epa0316.pdf>
- Instituto Nacional de Seguridad e Higiene en el Trabajo, INSHT. (2007). *VI Encuesta Nacional de Condiciones de Trabajo (ENCT VI)*. Barcelona: INSHT.
- Instituto Nacional de Seguridad e Higiene en el Trabajo. (2011). *VII Encuesta Nacional de Condiciones de Trabajo (ENCT VII)*. Madrid: INSHT.
- Instituto Nacional de Seguridad e Higiene en el Trabajo, INSHT. (2009). *Encuesta Nacional de Gestión de la Seguridad y Salud en las Empresas*. Barcelona: INSHT.
- Instituto Nacional de Seguridad e Higiene en el Trabajo (2015). *Algunas orientaciones para evaluar los factores de riesgo psicosocial* (ed. ampliada 2015). Madrid: Instituto Nacional de Seguridad e Higiene en el Trabajo. Recuperado de <http://www.insht.es/InshtWeb/Contenidos/Documentacion/FICHAS%20DE%20PUBLICACIONES/EN%20CATALOGO/PSICOSOCIOLOGIA/Maqueta%2018%204%20Angel%20lara.pdf>
- Instituto de Salud Pública y Laboral de Navarra (2015) *INSL. Factores Psicosociales. Identificación de situaciones de riesgo. DESCATALOGADO*. Nota informativa. Navarra: ISPLN. Recuperado de <http://www.navarra.es/NR/rdonlyres/08E25EAE-3183-4AA9-80E6-C14AE79D1385/321350/Notaconfirma.pdf>

- Instituto Sindical De Ambiente, Trabajo y Salud [ISTAS] (2010). *Manual del método CoPsoQistas21 (versión 1.5) para la evaluación y prevención de los riesgos psicosociales para empresas con 25 o más trabajadores y trabajadoras*. Edit. Instituto Sindical de Trabajo, Ambiente y Salud. Barcelona. Recuperado de <http://www.istas.net/copsoq/>
- Inspección de Trabajo y Seguridad Social (ITSS) (2012). *Guía de Actuaciones de la Inspección de Trabajo y Seguridad Social sobre Riesgos Psicosociales*. Recuperado de http://www.laboral-social.com/files-laboral/Guia_psicosociales.pdf
- Inspección de Trabajo y Seguridad Social (ITSS) (2014). *Informe Anual de la Inspección de Trabajo y Seguridad Social 2014*. Madrid: Ministerio de Empleo y Seguridad Social. Recuperado http://www.empleo.gob.es/itss/ITSS/ITSS_Descargas/Que_hacemos/Memorias/Memoria_2014_2.pdf
- Inspección de Trabajo y Seguridad Social (ITSS) (2015). *Informe Anual de la Inspección de Trabajo y Seguridad Social 2015*. Madrid: Ministerio de Empleo y Seguridad Social. Recuperado http://www.empleo.gob.es/itss/ITSS/ITSS_Descargas/Que_hacemos/Memorias/Memoria_2015_w.pdf
- Instituto Sindical de Ambiente, Trabajo y Salud. ISTAS. (2010). *Manual del método CoPsoQ - istas21 (versión 1.5) para la evaluación y prevención de los riesgos psicosociales para empresas con 25 o más trabajadores y trabajadoras*. Barcelona: Instituto Sindical de Trabajo, Ambiente y Salud.
- Ivancevich, J. M. (2001). La ambigüedad sobre el futuro en el empleo. En S.L. Sauter, L.R. Murphy, J.J. Hurrell y L. Levi (Dirs.). *Factores psicosociales y de organización. Enciclopedia de Salud y Seguridad en el Trabajo* (cap.34, pp.34-35). Madrid: Organización Internacional del Trabajo.
- Ivancevich, J. M. y Matteson, M. T. (1980). *Stress and Work*. Scott Foresman, Glenview, Illinois.
- Ivancevich, J. M. y Matteson, M. T. (1983). *Stress Diagnostic Survey*, University of Houston, Houston, TX.
- Jackson, D. N. (1986). The proces of responding in personality assessment. En A. Angleitner & J.S. wiggins (Eds.). *Personality assessment via questionnaires* (pp.123-143). Berlin: Springer-Verlag.

- Jackson D.N. y Messick S. (1958). Content and style in personality assessment. *Psychological Bulletin*, 55, 243-252.
- Jackson, D.N. y Messick, S.J. (1961). Acquiescence and desirability as response determinants on the MMPI. *Educational and Psychological Measurement*, 4, 771-790.
- Jacobson, L. I., Kellogg, R. W., Cauce, A. y Slavin, R. S. (1977). A multidimensional social desirability inventory. *Bulletin of the Psychonomic Society*, 9(2), 109-110.
- Jansen, A., König, C. J., Kleinmann, M. y Melchers, K. G. (2012). The interactive effect of impression motivation and cognitive schema on self-presentation in a personality inventory. *Journal of Applied Social Psychology*, 42(8), 1932-1957. doi:10.1111/j.1559-1816.2012.00925.x
- Jáñez, L. (1989). *Fundamentos de Psicología Matemática*. Madrid: Pirámide.
- Jenkins, G. (1984). *Response sets and personality measures: The K scale of the MMPI*. Texas Tech University, Lubbock.
- Jensen, B. J. (1995). Los cuestionarios de autoinforme en la evaluación conductual. En G. Buela-Casal, V. E. Caballo y J. C. Sierra (dirs.), *Manual de evaluación en psicología clínica y de la salud* (pp. 109-130) Madrid: Siglo XXI.
- Jex, S. M. (2001). Claridad y sobrecarga de los roles asignados. En S.L. Sauter, L.R. Murphy, J.J. Hurrell y L. Levi (Dirs.). *Factores psicosociales y de organización. Enciclopedia de Salud y Seguridad en el Trabajo* (cap. 34, pp. 31-32) Madrid: Organización Internacional del Trabajo.
- Jex, S. M. y Crossley, C. D. (2005). Organizational consequences. En J. Barling, E. K. Kelloway y M. R. Frone (Eds.), *Handbook of work stress* (pp. 575-599). Thousand Oaks: Sage Publications.
- Jiménez, S. (2014). Factores psicosociales y salud: de vida laboral a factor de riesgo. *Revista Española de Comunicación en Salud*, 5(2), 134-148.
- Jiménez, F. y Sánchez, G. (2004). La falsificación de las técnicas psicométricas: un estudio con el MMPI - 2. *I Congreso de Psicología Jurídica y Forense en Red*. Colegio Oficial de Psicólogos Madrid.

- Jodoin, M. G. y Gierl, M. J. (Ed) (2001). Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14(4), 329-349.
- Johannson, G., Johnson, J. V. y Hall, E. M. (1991). Smoking and sedentary behavior as related to work organization. *Social Science and Medicine*, 32, 837-846.
- Johns, G. y Saks, A. M. (2005). *Organizational Behaviour* (6th ed.). Toronto: Pearson Education Canada Inc.
- Johnson, J.V. y Hall, B.M. (1988). Job strain, work place social support and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health*, 78(10), 1338-1342.
- Johnson, J.V. y Hall, B.M. (1994). Social support in the work environment and cardiovascular disease. En S. Shumaler y S. Czajkowski. (Eds.), *Social Support and Cardiovascular Disease*. Nueva York: Baywood.
- Jones, E. y Sigall, H. (1971). The Bogus Pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76(5), 349-364.
- Jones, F. y Fletcher, B. C. (2003). Job control, physical health and psychological well-being. En M. J. Schabracq, J. A. M. Winnubst y C. L. Cooper (Eds.), *The Handbook of Work and Health Psychology* (pp. 383-425). Chichester: Wiley and Sons.
- Jones, F. y Kinman, G. (2002). Approaches to studying stress. En F. Jones y J. Bright (eds.) *Stress: Myth, Theory and Research*. 2a ed. (pp. 17-45). Londres: Prentice Hall.
- Jones, J. W. (1980). *Staff Burnout Scale for Health Professionals* (SBS-HP). Park Ridge, Illinois: London House.
- Jöreskog, K. G. (1977). Factor analysis by least-squares and maximum-likelihood methods. En K. Enslein, A. Ralston y H.S. Wilf (Eds.), *Statistical methods for digital computers*, Vol. 3. New York: Wiley.
- Judge, T.A., Heller, D. y Mount, M.K. (2002). Five-factor model of personality and job satisfaction. A meta-analysis. *Journal of Applied Psychology*, 87(3), 530-541.

- Justicia, F., Fernández, E., Benítez, J. L., Villena, M. D., García, A. y Caurcel, M. J. (2002). *Cuestionario sobre Acoso Laboral en Contextos Universitarios*. Unpublished Scientific Paper. Registro nº GR-1196-02.
- Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D. y Rosenthal, R. A. (1964). *Organizational stress. Studies in role conflict and ambiguity*. Nueva York: Wiley and Son.
- Kaiser, H.F. (1970). A second generation little jiffy. *Psychometrika*, 35, 401-415.
- Kaiser, H. F. y Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34, 111-117.
- Kalimo, R. (1988). Los factores psicosociales y la salud de los trabajadores: panorama general. En Kalimo R., El-Batawi M, Cooper C. L. (comps.) *Los factores psicosociales en el trabajo y su relación con la salud*. (pp. 3-8). OMS. Ginebra.
- Kamarck, T. W., Shiffman, S., Sutton-Tyrrell, K., Muldoon, M. F. y Tepper, P. (2012). Daily psychological demands are associated with 6-year progression of carotidartery atherosclerosis: The Pittsburgh Healthy Heart Project. *Psychosomatic Medicine*, 74, 432-439.
- Kan, M. Y. (2008). Measuring Housework Participation: the gap between "stylised" questionnaire estimates and diary-based estimates'. *Social Indicators Research*, 86(3), 381-400.
- Kaplan, A., Skogstad, L. y Girschick, M. A. (1949). *The prediction of social and technological events*. RAND Corporation
- Karasek, R. (1979). Job demands, job decision latitude, and mental strain: implications for job redesign. *Administrative Science Quarterly*, 24, 285-309.
- Karasek, R. A., Brisson, C., Kawakami, N., Houtman, I., Bongers, P. y Amick, B. (1998). The job content questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *The Journal of Occupational Health Psychology*, 3(4), 322-355.
- Karasek, R. A., Schwartz, J. y Theorell, T. (1982). *Job characteristics, occupation, and coronary heart disease*. (Final report on Grant No. R-01-OH00906). Cincinnati OH: National Institute for Occupational Safety and Health.

- Karasek, R. A. y Theorell, T (1990). *Healthy work. Stress, productivity, and the reconstruction of working life*. Nueva York: Basic Books.
- Karasek, R., Triantis, K. y Chaudhry, S. (1982). Co-worker and Supervisor Support as Moderators of Associations Between Task Characteristics and Mental Strain. *Journal of Occupational Behaviour*, 3, 147-160.
- Kasl, S. V. (1992). Surveillance of psychological disorders in the workplace. En G. P. Keita y S. L. Sauter (eds.) *Work and WelBeing: An Agenda for the 1990s*. American Psychological Association, Washington D.C.
- Kauko, K. y Palmroos, P. (2014). The Delphi Method in Forecasting Financial Markets – An Experimental Study. *International Journal of Forecasting*, 30, 313-327.
- Keashly, L. (1998). Emotional abuse in the workplace: Conceptual and empirical issues. *Journal of Emotional Abuse*, 1, 85-117.
- Keinan, G. (1987). Decision-making under stress: Scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology*, 52, 639-644.
- Kelly, G. (1955). *The Psychology of Personal Constructs* (2nd vol.). New York: Norton.
- Kendler, K.S., Kessler, R.C., Neale, M.C., Heath, A.C. y Eaves, L.J. (1993). The prediction of major depression in women: Toward an integrated etiologic model. *American Journal of Psychiatry*, 150, 1139–1148.
- Kent, M (2003). *Diccionario Oxford de Medicina y Ciencias del Deporte* (trad. De P.G. del Campo. Rev. de M. Rius). Paidotribo, Barcelona.
- Kerlinger, F. (1975). *Foundations of Behavioral Research*. New York: Holt, Rinehartand Winston, Inc.
- Kerlinger, F. N. y Lee, H. B. (2002). *Investigación del Comportamiento. Métodos de Investigación en Ciencias Sociales* (4a ed.). México: Mc Graw-Hill.
- Kerr, J. H. (1996). Employee Fitness Programmes and Reduced Absenteeism: A Case Study. En Kerr, J. Griffiths, A. & Cox, T. (Eds.), *Workplace Health - employee fitness and exercise* (Chapter 10, 159-167). London, London.

- Kerr, M.S. (1998). *Workplace Psychosocial Factors and Musculoskeletal Disorders: A Discussion Paper*. Institute for Work & Health, Toronto, Ontario, pp. 7-16.
- Kiecolt-Glaser, J. K., Cacioppo, J. T., Malarkey, W. B. y Glaser, R. (1992). Acute psychological stressors and short-term immune changes: What, why, for whom, and to what extent? *Psychosomatic Medicine*, 54, 680-685.
- Kiecolt-Glaser, J. K., Marucha, P. T., Malarkey, W. B., Mercado, A. M. y Glaser, R. (1996). El estrés psicológico enlentece la cicatrización de las heridas. *Lancet* 28, 188-191.
- Kim, J. J. y Diamond, D. M. (2002). The stressed hippocampus, synaptic plasticity and lost memories. *Nature Reviews: Neuroscience*, 3, 453-462.
- Kim, E. J. y Dimsdale, J. E. (2007). The effect of psychosocial stress on sleep: a review of polysomnographic evidence. *Behavioral Sleep Medicine*, 5, 256-278.
- King, M. F. y Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology y Marketing*, 17(2), 79-103.
- Kishchuk, N., Peters, C., Towers, A.M., Sylvestre, M., Bourgault, C. y Richard, L. (1994). Formative and effectiveness evaluation of a worksite program promoting healthy alcohol consumption. *Am J Health Promot*, 8, 353-362.
- Kivimäki, M., Elovainio, M., Vahtera, J., Virtanen, M. y Stansfeld, S. A. (2003). Association between organizational inequity and incidence of psychiatric disorders in female employees. *Psychological Medicine*, 33, 319-326.
- Kivimäki, M., Leino-Arjas, P., Luukkonen, R., Riihimäki, H., Vahtera, J. y Kirjonen, J. (2002). Work stress and risk of cardiovascular mortality: Prospective cohort study of industrial employees. *British Medical Journal*, 325(7369), 857-863.
- Kivimäki, M., Viortanen, M., Elovainio, M., Louvonen, A., Vaananen, A. y Vahtera, J. (2006). Work stress in the etiology of coronary heart disease: A meta-analysis. *Scandinavian Journal of Work Environment and Health*, 32, 431-432.
- Knowles, E. S. y Nathan, K. (1997). Acquiescent responding in self reports: Cognitive style or social concern? *Journal of Research in Personality*, 31, 293-301.

- Kohan, A. y Mazmanian, D. (2003). Police work, bournout, and pro-organizational behavior. A consideration of daily work experiences. *Criminal Justice & Behavior*, 30, 559-583.
- Kompier, M. A. J. (2002). Job design and well-being. En M. Schabracq, J. Winnubst y C. Cooper (Eds.), *Handbook of work and health psychology* (pp. 429-454). Chichester, England: Wiley.
- Konrad, A.M. y Linnehan, F. (1995). Formalized HRM Structures: Coordinating Equal Employment Opportunity or Concealing Organizational Practice? *Academy of Management Journal*, 38, 787-820.
- Kouvonen, A., Kivimäki, M., Virtanen, M., Pentti, J. y Vahtera, J. (2005). Work stress, smoking status, and smoking intensity: An observational study of 46,190 employees. *Journal of Epidemiology and Community Health*, 59, 63-69.
- Krenz, C. y Sax, G. (1987). Acquiescence as a function of test type and subject uncertainty. *Educational Psychologica Measurement*, 47, 575-581.
- Kreuter, F., Presser, S. y Tourangeau, R. (2008). Social desirability bias in CATI, IVR and Web Surveys. *Public Opinion Quartely*, 72(5), 847-865.
- Kristensen, T.S. (1996). Job stress and cardiovascular disease: a theoretic critical review. *Journal of Occupational Health Psychology*, 1(3), 246-260.
- Kristensen, T. S. (2010). A questionnaire is more than a questionnaire. *Scandinavian Journal of Public Health*, 38(3), 149-155.
- Kristensen, T. S., Borritz, M., Villadsen, E. y Christensen, K.B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work and Stress*, 19(3), 192-207. doi: 10.1080/02678370500297720
- Kristensen, T. S., Hannerz, H., Høgh, A. y Borg, V. (2005). The Copenhagen Psychosocial Questionnaire – A tool for the assessment and improvement of the psychosocial work environment. *Scandinavian Journal of Work, Environment and Health*, 31(6), 438-449.
- Krosnick, J. A., Narayan, S. y Smith, W. R. (1996). Satisficing in surveys: Initial evidence. En M. T. Braverman y J. K. Slater (Eds.), *Advances in Survey Research* (pp. 29-44). San Francisco: Sage.

- Kulas, J. T. y Stachowski, A. A. (2009). Construct and response bias correlates in summated scale definitions of personality traits. En F. Columbus (Ed.), *Personality Assessment: New Research*. Hauppauge, NY: Nova Science Publishers.
- Kuper, H. y Marmot, M. (2003). Job strain, job demands, decision latitude, and risk of coronary heart disease within the Whitehall II study. *Journal of Epidemiology & Community Health*, 57(2), 147–153.
- Labrador, F. J. (1992). *El estrés: Nuevas Técnicas para su control*. Madrid: Temas de Hoy.
- Labrador, F. J. y Crespo, M. (1993). *Estrés: trastornos psicofisiológicos*. Madrid: Eudema.
- Lahera, M. (2006). Método del Instituto Navarro de Salud Laboral. En J. L. Meliá, C. Nogareda, M. Lahera, A. Duro, J. M. Peiró, R. Pou y otros, *Perspectivas de intervención en Riesgos Psicosociales. Evaluación de Riesgos* (pp. 64-82). Barcelona: Foment del Treball Nacional.
- Lahera, M. y Góngora, J. J (2002). *Factores psicosociales. Identificación de situaciones de riesgo*. Navarra: Instituto Navarro de Salud Laboral.
- Lahera, M. y Nogareda, C. (2009). El método del INSL para la identificación y evaluación de factores psicosociales. INSHT (NTP: 840). Recuperado de <http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/821a921/840%20web%20.pdf>
- Lajunen, T. y Summala, H. (2003). Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. *Transportation Research Part F*, 6, 97-107.
- Lambert, E. G., Hogan, N. L., Camp, S. D. y Ventura L. A. (2006). The impact of work–family conflict on correctional staff: A preliminary study. *Criminology and Criminal Justice*, 6, 371–387.
- LaMontagne, A. D., Keegel, T., Louie, A. M. y Ostry, A. (2010). Job stress as a preventable upstream determinant of common mental disorders: A review for practitioners and policy-makers. *Advances in Mental Health*, 9, 17-35.
- Landeta, J. (2002). *El método Delphi. Una técnica de previsión del futuro*. Ariel. Barcelona

- Landeta, J. (2006). Current Validity of the Delphi Method in Social Sciences. *Technological Forecasting and Social Change*, 73, 467-482
- Landeta, J. y Barrutia, J. (2011). People consultation to construct the future: a Delphi application. *International Journal of Forecasting*, 27(1), 134-151.
- Landeta, I., Mate, I., Ruiz, V. y Galter, J. (2008). Results of a Delphi survey in drawing up the input – output tables for Catalonia. *Technological Forecasting and Social Change*, 75, 32-56.
- Landis J. R. y Koch G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Landsbergis, P. A., Schnall, P. L., Belkic, K. L., Baker, D., Schwartz, J. y Pickering, T. G. (2001). Work stressors and cardiovascular disease. *Work*, 17, 191-208.
- Landsbergis, P. A., Schnall, P. L., Belkic, K. L., Baker, D., Schwartz, J. E. y Pickering, T. G. (2003). The workplace and cardiovascular disease: Relevance and potential role for occupational health psychology. En J. C. Quick & L. E. Tetrick (Eds.), *Handbook of occupational health psychology* (pp. 265–287). Washington, DC: American Psychological Association.
- Lang, T. (2000). An overview of four futures methodologies (Delphi, environmental scanning, issues management and emerging issue analysis).” *The Manoa Journal of Fried and Half-Fried Ideas (about the future)* (Online). Hawaii Research Center for Futures Studies, University of Hawaii at Monoa, Honolulu, HI. Recuperado de <http://www.futures.hawaii.edu/publications/half-fried-ideas/J7/LANG.pdf>
- Larsen, R. (2001). Decision-Making by Military Students under Extreme Stress. *Military Psychology*, 13(2), 89-92.
- Larsman, P. (2006). *On the relation between psychosocial work environment and musculoskeletal symptoms*, Stockholm: Department of Psychology, Goteborg University.
- Lattin, J., Carroll, D.J. y Green, P.E. (2003). *Analyzing multivariate data*. Pacific Grove, CA: Duxbury Press.

- Lauridsen, O. y Tonnesen, T. (1990). Injuries related to aspects of shiftworking. A comparison of different offshore shift arrangements. *Journal of Occupational Accidents*, 12, 167-176.
- Lazarus, R.S. y Folkman, S. (1986). *Estrés y procesos cognitivos*. Barcelona: Martínez Roca.
- Leka, S., Griffinths, A. y Cox, T. (2004). *La organización del trabajo y el estrés. Serie Protección de la salud de los trabajadores n°3*. Recuperado de http://www.who.int/occupational_health/publications/en/pwh3sp.pdf
- Leka, S. y Jain, A. (2010). *Health Impact of Psychosocial hazards at work: An overview*. Brussels: WHO. Recuperado de http://apps.who.int/iris/bitstream/10665/44428/1/9789241500272_eng.pdf
- Lemos, V. (2005). Construcción y validación de una Escala para la Evaluación de la Deseabilidad Social Infantil (EDESI). *Interdisciplinaria*, 22(1), 79–96.
- Lentz, T. F. (1938). Acquiescence as a factor in the measurement of personality. *Psychological Bulletin*, 35, 659.
- León, O. G. y Montero, I. (2003). *Métodos de investigación en psicología y educación* (3ª ed.) Madrid: McGraw-Hil.
- Lester, N, Nebel, L. E. y Baum, A. (1994). Psychophysiological and behavioral measurement of stress. Applications to mental health. En W.R. Avison & I.H. Gotlib (Eds), *Stress and Mental Health. Contemporary issues and prospects for the future*. Nueva York: Plenum Press.
- Levenstein, S. (1998). Stress and peptic ulcer: Life beyond Helicobacter. *British Medical Journal*, 316, 538–541.
- Levenstein, S. (2002). Psychosocial Factors in Peptic Ulcer and Inflammatory Bowel Disease. *Journal of Consulting and Clinical Psychology*, 70(3), 739-750. doi: 10.1037//0022-006X.70.3.739
- Levenstein, S., Prantera, C., Varvo, V., Arcà, M., Scribano, M. L., Spinella, S. y Berto, E. (1996). Long-term symptom patterns in duodenal ulcer: Psychosocial factors. *Journal of Psychosomatic Research*, 41, 465-472.

- Levenstein, S., Prantera, C., Varvo, V., Scribano, M. L., Andreoli, A., Luzi, C., ... Marcheggiano, A. (2000). Stress and exacerbation in ulcerative colitis: A prospective study of patients enrolled in remission. *American Journal of Gastroenterology*, 95, 1213–1220.
- Levi, L. (2001). Factores psicosociales, estrés y salud. En S. L. Sauter, L. R. Murphy, J. J. Hurrell y L. Levi (Dirs.), *Factores psicosociales y de organización. Enciclopedia de Salud y Seguridad en el Trabajo* (cap.34, pp. 3-6) Madrid: Organización Internacional del Trabajo.
- Levi, L., Frandenhaeuser, M. y Gardell, B. (1986). The characteristics of the workplace and the nature of its social demands. En S. G. Wolf y A. J. Finestone (Dires.): *Occupational Stress: Health and Performance at Work*. Littleton Mass: PSG.
- Levin, I. y Stokes, J. P. (1989). Dispositional approach to job satisfaction: role of negative affectivity. *Journal of Applied Psychology*, 74, 752-58.
- Ley 31/1995, de 8 de noviembre, de Prevención de Riesgos Laborales. *Boletín Oficial del Estado*. Madrid 10 de noviembre de 1995, núm 269, pp. 32590-32611.
- Leymann, H. (1986). *Vuxenmobbing: om psykiskt våld i arbetslivet* [Mobbing-psychological violence at work]. Lund: Studentlitteratur.
- Leymann, H. (1990a). Mobbing and psychological terror at workplaces. *Violence and Victims*, 5, 119-126.
- Leymann, H. (1990b). *Handbook for anvindning av LIPTformuläret för kartläggning av risker för psykiskt våld i arbetsmiljön* [Manual of the LIPT questionnaire for assessing the risk of psychological violence at work]. Estocolmo, Suecia: Violen.
- Leymann, H. (1992). *Leymann inventory of psychological terror*. Violen: Karlskrona.
- Leymann, H (1996a). *Mobbing. La persecution au travail*. Éd. Du Seuil, París.
- Leymann, H. (1996b). The content and development of mobbing at work. *European. Journal of Work and Organizational Psychology*, 5(2), 165-184.

- Leymann, H. (2009). *El contenido y el desarrollo del mobbing en el trabajo* (Sergio Navarrete, trad., obra original publicada en 1996). Canada: Universidad Waterloo. Recuperado de <http://www.kwesthues.com/e-LeymannEJWOP1996.pdf>
- Li, A. y Bagger, J. (2007). The balance inventory of desirable responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement*, 67(3), 525-538. doi:10.1177/0013164406292087
- Likert, R. (1932). A Technique for the Measurement of Attitudes. *Achieves of Psychology*, 140, 44-53.
- Lindström, K., Elo A-L, Skogstad, A., Dallner, M., Gamberale, F., Hottinen... Ørhede, E. (2000). *User's Guide for the QPSNordic, TemaNord 2000:603, General Nordic Questionnaire for Psychological and Social Factors at Work*. Nordic Council of Ministers: Copenhagen.
- Linstone, H. A. y Turoff, M. (1975). Introduction. En H. A. Linstoney M. Turoff (Eds.). *The Delphi method: Techniques and applications* (pp. 3-12). Reading, MA: Addison-Wesley Publishing Company
- Lipton, M. A. (1976). Behavioral effects of hypothalamic polypeptide hormones in animals and man. En E. J. Sachar (ed.), *Hormones, Behaviour and Psychopathology*. Raven, New York.
- Littlewood, H. (2006). Varianza común del método: el caso de los cuestionarios. *Revista Interamericana de Psicología Ocupacional*, 25(2), 57-70.
- Llaneza, F. J. (2009). *Ergonomía y psicología aplicada. Manual para la formación del especialista* (12ª ed.). Valladolid: Lex Nova.
- Llaneza, F. J. (2013). *La ergonomía forense y el papel de los ergónomos como peritos judiciales*. (tesis doctoral). Universidad de Oviedo.
- Llorens, S., del Líbano, M. y Salanova, M. (2009). Modelos teóricos de salud ocupacional. En M. Salanova (Ed.), *Psicología de la Salud Ocupacional*. (pp. 63-93). Madrid: Síntesis.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. doi: 10.6018/analesps.30.3.199361

- López-Cabarcos, M. y Vázquez-Rodríguez, P. (2005). *Mobbing: cómo prevenir, identificar y solucionar el acoso psicológico en el trabajo*. Madrid: Pirámide.
- Lorenz, K. (1963). Das sogenannte Böse: Zur Naturgeschichte der Aggression. Vienna: Borotha-Schoeler. [Trad. esp.: *Sobre la agresión: el pretendido mal*. Madrid: Siglo XXI, 1992]
- Lorenz, K. (1965). *Evolution and Modification of Behaviour*: Chicago; University of Chicago Press.
- Lorenzo-Seva, U. (1999). Promin: A method for oblique factor rotation. *Multivariate Behavioral Research*, 34, 347-356
- Lorenzo-Seva, U. (2003). A factor simplicity index. *Psychometrika*, 68(1), 49-60.
- Lorenzo-Seva, U. y Ferrando, P. J. (2009). Acquiescent responding in partially balanced multidimensional scales. *British Journal of Mathematical and Statistical Psychology*, 62 (2), 319-326.
- Lorenzo-Seva, U. y Ferrando, P. J. (2010). Acquiescence as a source of bias and model and person misfit: A theoretical and empirical analysis. *British Journal of Mathematical Statistical Psychology*, 63(2), 427-448.
- Lorenzo-Seva, U. y Ferrando, P. J. (2013). FACTOR 9.2: A comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models. *Applied Psychological Measurement*, 37(6), 497-498. doi: 10.1177/0146621613487794
- Lorenzo-Seva, U. y Ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology, European Journal of Research Methods for the Behavioral and Social Sciences*, 2, 57-64.
- Louzán, R. (2014). Exposición a los factores psicosociales adversos en una muestra de trabajadores de la construcción de la provincia de Lugo. *Seguridad y Salud en el Trabajo*, 77, 28-35.
- Luceño, L. (2005). *Evaluación de factores psicosociales en el entorno laboral. Construcción y validación del Cuestionario Multidimensional DECORE*. Tesis doctoral. Madrid: Editorial Complutense.

- Luceño, L. y Martín, J. (2005). Estrés laboral: factores estresantes y adaptación. En J. L. Arco Tirado (ed.), *Estrés y trabajo: cómo hacerlos compatibles* (pp. 45-65). Sevilla: Instituto Andaluz de Administraciones Públicas.
- Luceño, L., Martín, J., Díaz, E. M. y Rubio, S. (2008). Un instrumento de evaluación de riesgos psicosociales en el entorno laboral. El cuestionario DECORE. *EduPsykhé*, 7(2), 131-153. Recuperado de http://dialnet.unirioja.es/servlet/dfichero_articulo?codigo=2800125
- Luceño, L., Martín, J., Jaén, M. y Díaz, E. M. (2005). Evaluación de Factores Psicosociales en el Entorno Laboral, *EduPsykhé*, 3(1), 19-41.
- Luceño, L., Martín, J., Rubio, S. y Jaén, M (2008). Psicología y Riesgos Laborales Emergentes, los Riesgos Psicosociales. *EduPsykhé*, 7(2), 111-129.
- Luceño, L., Martín, J., Rubio, S. y Díaz, E. M. (2010). Análisis Factorial Confirmatorio Del Cuestionario Decore. *Ansiedad y Estrés*, 16(2-3), 237-248.
- Luceño, L. M., Martín, J., Tobal, J. J. y Jaén, M. (2005). El Cuestionario Multidimensional DECORE: Un instrumento para la evaluación de factores psicosociales en el entorno laboral. *Ansiedad y Estrés*, 5, 189-202.
- Lucini, D., Silvano Riva, S., Pizzinelli, P. y Pagani, M. (2007). Stress management at the worksite. Reversal of symptoms profile and cardiovascular dys regulation. *Hypertension* 49, 291-297.
- Ludwig, B. (1997). Predicting the future: Have you considered using the Delphi methodology? *Journal of Extension*, 35(5), 1- 4. Recuperado de <http://www.joe.org/joe/1997october/tt2.html>
- Macht, M. (2008). How emotions affect eating: a five-way model. *Appetite*, 50(1), 1-11.
- Maciá, A. (1984). Concepto de psicología matemática. *L'Arrel*. 5, 107-126.
- Magnus, K., Diener, E., Fujita, F. y Pavot, W. (1993). Extraversion and neuroticism as predictors of objective life events: A longitudinal analysis. *Journal of Personality and Social Psychology*, 65, 1046-1053.

- Mahmood, M., Coons, S. J., Guy M. C. y Pelletier K. R. (2010). Development and testing of the Workplace Stressors Assessment Questionnaire. *Journal of occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 52(12), 1192-200. doi:10.1097/JOM.0b013e3181fb53dc.
- Man, T.S. (2013). Do personality and self-construal predict response style in self-rating scales? *Discovery-SS Student E-Journal* 29(2), 29-48. Recuperado de <http://ssweb.cityu.edu.hk/download/RS/E-Journal/Vol2/journal3.pdf>
- Mandler, G. (1982). The structure of value: accounting for taste. En: M.S. Clark y S.T. Fiske (eds.), *Affect and cognition*. Hillsdale, NJ: Lawrence Erlbaum.
- Mangione, T. W. y Quinn, R. P. (1975). Job Satisfaction, Counterproductive Behaviour and Drug Use at Work. *Journal of Applied Psychology*, 60, 114-116.
- Manning, M. R. y Osland, J. S. (1989). The relationship between absenteeism and stress. *Work & Stress*, 3(3), 223-235.
- Manoliadis, O. y Tsolas, I. (2010). Sustainability as an innovation in public procurement. The case of sustainable construction in Greece (2010), *Proceedings IPPC4 4th International Public Procurement Conference*, Seoul, South Korea, August 26-28, pp. 1-17.
- Mansilla, F. y Favieres, A. (2011). *Factores psicosociales en el trabajo*. Madrid Salud. Recuperado de <http://www.madridsalud.es/publicaciones/saludpublica/RiesgosPSICOSO CIALES.pdf>
- Marcos, J. I. y Velázquez M. (2011). Observatorio Vasco sobre Acoso Moral en el Trabajo/Basque Workplace Mobbing Observatory. *Medicina Social*, 6(4), 354-357.
- Marín, M. (2013). Adicciones que dificultan las tareas profesionales. *Trabajo Social Hoy*, 69, 81-96. doi:10.12960/TSH.2013.0012
- Marquis, K. H., Marquis, M. S. y Polich, J. M. (1986). Response bias and reliability in sensitive topic surveys. *Journal of the American Statistical Association*, 81, 381-389.
- Marmot, M., Bosma, H., Hemingway, H., Brunner, E. y Stansfeld, S. (1997). Contribution of job control and other risk factors to social variations in coronary heart disease incidence. *Lancet*, 350, 235-239.

- Marmot, M., Feeney, A., Shipley, M., North, F. y Syme, L. (1995). Sickness absence as a measure of health status and functioning: from the UK Whitehall II study. *Journal Epidemiology Community Health*, 49, 124-130.
- Marsh, H. W., Hau, K. T. y Grayson, D. (2005). Goodness of fit in structural equation models. En A. Maydeu-Olivares y J. J. McArdle (Eds.), *Contemporary Psychometrics: A festschrift* (pp. 225-340). Mahwah, NJ: Lawrence Erlbaum Associates.
- Martin, J. (1964). Acquiescence: Measurement and theory. *British Journal of Social and Clinical Psychology*, 3, 216-225.
- Martín, J., Luceño, L., Jaén, M. y Rubio, S., (2007). Relación entre factores psicosociales adversos, evaluados a través del Cuestionario Multidimensional DECORE y salud laboral deficiente. *Psicothema*, 19(1), 95-101.
- Martin, R. J. y Hine, D. W. (2005). The development and validation of the Uncivil Workplace Behavior Questionnaire. *Journal of Occupational Health Psychology*, 10, 477-490.
- Martín, P., Salanova, M. y Peiró, J.M. (2003). El estrés laboral. ¿Un concepto cajón-de-sastre?. *Revista de Relaciones Laborales y Ciencias del Trabajo*, 10, 167-185.
- Martín-Daza, F. (1993). *El estrés: proceso de generación en el ámbito laboral. (NTP. 318)*. Instituto Nacional de Seguridad e Higiene en el Trabajo (INSHT). Recuperado de http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/301a400/ntp_318.pdf
- Martín-Daza, F. y Pérez-Bilbao, J. (1997). *Factores psicosociales: metodología de evaluación. (NTP 443)*. Barcelona: Instituto Nacional de Seguridad e Higiene en el Trabajo (INSHT).
- Martín-Daza, F., Pérez-Bilbao, J. y López, A. (1998). *El hostigamiento psicológico en el trabajo: mobbing (NTP.476)*. Madrid: Instituto Nacional de Seguridad e Higiene en el Trabajo (INSHT).
- Martínez, E. (2003). La técnica delphi como estrategia de consulta a los implicados en la evaluación de programas. *Revista de Investigación Educativa*, 21, 2, 449-463

- Martínez, I. M. y Salanova, M. (2003). Niveles de burnout y engagement en estudiantes universitarios. Relación con el desempeño y desarrollo profesional. *Revista de Educación*, 330, 361-371.
- Martínez, M., Irurtia, M. J., Martínez, C., Torres, H. y Queipo, D. (2012). El acoso psicológico en el trabajo o mobbing: patología emergente. *Gaceta Internacional de Ciencias Forenses*, 3, 1-12.
- Martínez-Arias, M.R. (1995). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid: Síntesis.
- Martínez-Plaza, C.A. (2001). *Estrés: aspectos médicos*. Instituto Nacional de Seguridad e Higiene en el Trabajo. Madrid.
- Martínez-Plaza, C. A. (2007). Riesgos psicosociales. Repercusión neurológica del estrés. En: *Ictus y actividad laboral. Factores emergentes de riesgo ocupacional y aspectos socio sanitarios*. (pp. 1-14). Barcelona: Grupo de Estudio Neurología del Trabajo. http://neurologiadeltrabajo.sen.es/curso_nt_ictus_2007.htm
- Martínez-Plaza, C.A (2008). Estrés. Aspectos médicos. *Revista La Mutua*, 18, 67-92.
- Martínez-Plaza, C.A (2009a). Estrés laboral y trastornos musculoesqueléticos (I). *Gestión Práctica de Riesgos Laborales*, 61, 38-46.
- Martínez-Plaza, C.A (2009b). Estrés laboral y trastornos musculoesqueléticos (II). *Gestión Práctica de Riesgos Laborales*, 62, 18-25.
- Martínez-Plaza, C. A. (2010). El estrés laboral como factor de riesgo de suicidio (I), (II), (III). *Gestión Práctica de Riesgos Laborales*, 74, 12-21.
- Martínez Selva, J. (2005). *La psicología de la Mentira*. Barcelona: Paidós.
- Maslach, C. (2009). Comprendiendo el Burnout. *Ciencia y Trabajo*, 11(32), 37-43.
- Maslach, C. y Jackson, S. E. (1981a). *MBI: Maslach Burnout Inventory. Manual*. Palo Alto: University of California, Consulting Psychologists Press.
- Maslach, C. y Jackson, S. E. (1981b). The measurement of experienced burnout. *Journal of Occupational Behaviour*, 2, 99-113.

- Maslach, C., Jackson, S. E. y Leiter, M. P. (1996). *Maslach Burnout Inventory*. (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.
- Matesanz, A. (1997). *Evaluación estructurada de la personalidad*. Madrid: Pirámide.
- MATRIX (2013). *Economic analysis of workplace mental health promotion and mental disorder prevention programmes and of their potential contribution to EU health, social and economic policy objectives. Final Report. EU Health Programme (2008-2013)*. Recuperado de https://ec.europa.eu/health/sites/health/files/mental_health/docs/matrix_economic_analysis_mh_promotion_en.pdf
- Mattews, D. (1990). A comparison of burnout in selected occupational fields. *The Career Development Quarterly*, 38(3), 230-239.
- Mawdsley, J. E. y Rampton D. S (2006). The role of psychological stress in inflammatory bowel disease. *Neuroimmunomodulation*, 3(5-6), 327-336. doi:10.1016/j.psyneuen.2011.05.014
- Mayer, E. A. (2000a). Psychological stress and colitis. *Gut*, 46, 595-596.
- Mayer, E. A. (2000b). The neurobiology of stress and gastrointestinal disease. *Gut*, 47(6), 61-869 doi: 10.1136/gut.47.6.861
- Mayoral, S. y Espluga, J. (2010). Mobbing: ¿un problema de perfiles psicológicos o un problema de organización del trabajo? Dos estudios de caso. *Cuadernos de Relaciones Laborales*, 28(2), 233-255.
- McClendon, M. J. (1991). Acquiescence and recency response-order effect in interview surveys. *Sociological Methods and Research*, 20(1), 60-103.
- McCrae, R. y Costa, P. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51(6), 882-888.
- McGee, R. K. (1967). Response set in relation to personality: An orientation. En I. A. Berg (Ed.), *Response set in personality assessment* (pp. 1-31). Chicago: Aldine.
- McGrath, J. E. (1970). *Social and Psychological Factors in Stress*. Nueva York: Holt, Rinehart Winston.

- McGrath, R. E., Mitchell, M., Kim, B. H. y Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, 136(3), 450–470. doi: 10.1037/a0019216
- McEwen, B. (2002). *The End of Stress as we know it*. Washington: Dana Press.
- Meade, A.W., Watson, A.M. y Kroustalis, M. (2007). Assessing common methods bias in organizational research, 22nd Annual Meeting of the Society for Industrial and Organizational Psychology Proceedings, New York, April.
- Meehl, P. E. y Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology*, 30(5), 525-564. Doi: 10.1037/h0053634
- Meisenberg, G., Lawless, E., Lambert, E. y Newton, A. (2006). The social ecology of intelligence on a Caribbean island. *Mankind Quarterly*, 46, 395–433.
- Mejía, J. C., Laca, F. y Gondra J. M. (2009). Factores de personalidad, afectivos y sociodemográficos en la predicción del bienestar laboral de docentes. *Psicología y Salud*, 19(1), 121-132.
- Melchior, M., Caspi, A., Milne, B.J., Danese, A., Poulton, R. y Moffitt, T.E. (2007). Work stress precipitates depression and anxiety in young, working women and men. *Psychological Medicine*, 37(8), 1119–1129.
- Meliá, J. L. (1990). *La construcción de la Psicometría coma ciencia teórica y aplicada*. Valencia: Cristóbal Serrano.
- Meliá, J. L. (1998). Un modelo causal psicosocial de los accidentes laborales. *Anuario de Psicología*, 29, 25-43.
- Meliá, J. L. (2003a). *Batería de factores psicosociales de salud laboral: Instrumentos diagnósticos para la prevención de disfunciones y patologías laborales*. Valencia: Ed. Cristóbal Serrano. Recuperado de <http://www.uv.es/seguridadlaboral>
- Meliá, J. L. (2003b). *Bateria Valencia –PREVACC: Instrumentos diagnósticos para la prevención de accidentes laborales*. Valencia: Ed. Cristóbal Serrano. Recuperado de <http://www.uv.es/seguridadlaboral>

- Meliá, J. L. (2004a). La Batería de Factores Psicosociales de la Universidad de Valencia: La evaluación multidimensional comprehensiva de los riesgos psicosociales. *En Trabajo presentado en el Tercer Congreso Internacional de Prevención de Riesgos Laborales*. (pp. 1–13). Santiago de Compostela. Recuperado de <http://www.uv.es/~meliajl/Papers/2004BFPSLMelia.pdf>
- Meliá, J. L. (2004b). El Modelo Causal Psicosocial de los Accidentes Laborales de la Universidad de Valencia: Perspectiva y nuevos desarrollos. *En Ponencia presentada en el Tercer Congreso Internacional de Riesgos Laborales*. Santiago de Compostela. Recuperado de <http://www.uv.es/~meliajl/Papers/2004ModeloMelia.pdf>
- Meliá, J. L. (2006). La Batería Valencia PREVACC de la Universidad de Valencia. En Foment del Treball (Ed.), *Perspectives de Intervenció en Riesgos Psicosociales. Evaluación de Riesgos* (pp. 155-180). Barcelona: Foment del Treball Nacional. Pags. 155-180.
- Meliá, J. L. (2008). *¿Cómo Evaluar los Riesgos Psicosociales en la Empresa? Metodologías, Oportunidades y Tendencias*. Universidad de Valencia. Recuperado de www.uv.es/seguridadlaboral
- Meliá, J. L., Nogareda, N., Lahera, M., Duro, A., Peiró, J. M., Salanova, M. y Gracia, D. (2006). Principios comunes para la evaluación de los riesgos psicosociales en la empresa. En Foment del Treball (Ed.), *Perspectives de Intervenció en Riesgos Psicosociales. Evaluación de Riesgos* (pp. 13-36). Barcelona: Foment del Treball Nacional.
- Mellenbergh, G. J. (1982). Contingency table models for assesing item bias. *Journal of Educational Statistics*, 7, 105-118.
- Mellenbergh, G. J. (1995). Conceptual notes on models for discrete polytomous item responses. *Applied Psychological Measurement*, 19, 91-100.
- Mendoza, J. (2012). La teoría en la investigación científica: marco teórico, modelos y medición. En K.A.C. Sáez, F.J. Gorjón, M. Gonzalo y C.M. Díaz (Eds.) *Metodología para investigaciones de alto impacto en las ciencias sociales* (pp. 43-94). Madrid: Dykinson.
- Merino, S. O. (2013). *Guía para la Prevención del Estrés en la Empresa*. Madrid: FREMAP.

- Meseguer, M., Soler, M. I. y García-izquierdo, M. (2007). Los factores psicosociales de riesgo en el trabajo como predictores del mobbing. *Psicothema*, 19(2), 225–230.
- Messick, S. (1966). The Psychology of Acquiescence: An Interpretation of Research Evidence1. *ETS Research Bulletin Series*, 1, i-44. doi: 10.1002/j.2333-8504.1966.tb00357.x.
- Messick, S. (1967). The psychology of acquiescence: an interpretation of the research evidence. En I. A. Berg (Ed.), *Response set in personality assessment* (pp. 115-145). Chicago: Aldine.
- Messick, S. (1968). Response sets. In: D.L. Sills (Ed.), *International encyclopedia of the social sciences* (pp. 492-496). New York: Macmillan.
- Messick, S. (1991). Psychology and methodology of response styles. En R. E. Snow y D. E. Wiley (Eds.), *Improving inquiry in social science: A volume in honor of Lee J. Cronbach* (pp. 161–200). Hillsdale, NJ: Erlbaum.
- Michaelis, W. y Eysenck, H.J. (1971). The determination of personality inventory factor patterns and intercorrelations by changes in real-life motivation. *Journal of Genetical Psychology*, 118, 223-234.
- Mikesell, R. H., Calhoun, L. G. y Lottman, T. J. (1970). Instructional set and the Coopersmith Self-Esteem Inventory. *Psychological Reports*, 26, 317-318.
- Miller, K. (2005): *Communication Theories: Perspectives, Processes, and Contexts*. (2nd ed.). Nueva York: McGraw-Hill.
- Miller, T. R. y Spray, J.A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of Educational Measurement*, 30, 107-122.
- Millham, J. y Kellogg, R. W. (1980). Need for social approval: Impression management or self-deception?. *Journal of Research in Personality*, 14(4), 445-457. doi: 10.1016/0092-6566(80)90003-3
- Milliken, G. (1979). The Delphi method of technological forecasting. En P.A. Hanle (Ed.), *High technology on earth: studies in using aerospace systems and methods* (pp. 30-42). Smithsonian Institution Press, Washington, D.C: Smithsonian Institution Press.

- Miner, F.C. Jr. (1984) Group versus individual decision making: A investigation of performance measures, decision strategies, and process losses/gains. *Organizational Behaviour and Human Performance*, 33, 112-124.
- Ministerio de Trabajo y Asuntos Sociales (2010). *Encuesta de Calidad de Vida en el Trabajo (ECVT)*. Madrid: Subsecretaría de Trabajo e Inmigración. Disponible en <http://www.empleo.gob.es/estadisticas/ecvt/Ecvt2010/index.htm>
- Mingote, J. C., Moreno, B. y Gálvez, M. (2004). Desgaste profesional y salud de los profesionales médicos: revisión y propuestas de prevención. *Medicina Clínica*, 123(7), 265-270.
- Mirowsky, J. y Ross, C. E. (1991). Eliminating defense and agreement bias from measures of the sense of control: A 2x2 index. *Social Psychology Quarterly*, 54, 127-145.
- Mitchell, V. W. (1991). The Delphi technique: An exposition and application. *Technology, Analysis & Strategic Management*, 3(4), 333-358.
- Molina, C. (2010). La promoción de la investigación en riesgos psicosociales relacionados con el trabajo en España: una asignatura pendiente. *Revista de Prevención de Riesgos Psicosociales y Bienestar en el trabajo*, 1, 9-11.
- Molina, C. y García, M. (2010). Más allá de un “cuestionario”: de la “evaluación de riesgos” a la “intervención psicosocial. *Revista de prevención de riesgos psicosociales y bienestar en el trabajo*, 2, 7-10.
- Molina, J. M. (2010). Metanálisis: Relación entre factores psicosociales en el trabajo y absentismo laboral. *Medicina y Seguridad del Trabajo*, 56(220), 200-219.
- Moncada, S. (1998). Salud laboral. En F. Martínez Navarro, J.M. Antó, P.L. Castellanos, M. Gili, P. Marset y V. Navarro (Eds.). *Salud Pública*. Madrid: Mcgraw-Hill /Interamericana.
- Moncada, S. y Artazcoz, L. (2000). Factores psicosociales. En F.G. Benavides, C. Ruiz Frutos y A.M. (Eds.) García. *Salud laboral* (2ª ed.). Barcelona: Masson.
- Moncada, S. y Llorens, C. (2006). *El método COPSQ (ISTAS 21, PSQCAT 21) de evaluación de riesgos psicosociales (NTP 703)*. Madrid: INSHT.

- Moncada, S., Llorens, C., Andrés, R., Moreno, N. y Molinero, E. (2014). *Manual del método CoPsoQ-istas21 (versión 2) para la evaluación y la prevención de los riesgos psicosociales en empresas con 25 o más trabajadores y trabajadoras*. Disponible en: [http://www.istas.net/copsoq/ficheros/documentos/v2/manual%20Copsoq%202\(24-07-2014\).pdf](http://www.istas.net/copsoq/ficheros/documentos/v2/manual%20Copsoq%202(24-07-2014).pdf)
- Moncada, S., Llorens, C., Font, A., Galtés, A. y Navarro, A. (2008). Exposición a riesgos psicosociales entre la población asalariada en España (2004-2005): valores de referencia de las 21 dimensiones del cuestionario COPSOQ/ISTAS21. *Revista Española de Salud Pública* 82(6), 667-675.
- Moncada, S., Llorens C., Gimeno X. y Font, A. (2007). Exposición laboral a riesgos psicosociales en población asalariada española. En S. Moncada y C. Llorens (Eds.). *Organización del trabajo, factores psicosociales y salud. Experiencias de prevención*. Madrid: ISTAS.
- Moncada, S., Llorens, A. C. y Kristensen, T. S. (2002). Manual para la evaluación de riesgos psicosociales en el trabajo. Método ISTAS 21(COPSOQ). Barcelona. Paralelo Edición, S.A.
- Moncada, S., Llorens, A. C., Kristensen, T. S. y Vega S. (2006). *La metodología COPSOQ (ISTAS21, PSQCAT21) de evaluación de riesgos Psicosociales*. NTP 703. INSHT. Recuperado de http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnica/s/NTP/Ficheros/701a750/ntp_703.pdf
- Moncada, S., Pejtersen, J. H., Navarro, A, Llorens, C., Burr, H., Hasle, P y Bjorner, J. B. (2010). Psychosocial work environment and its association with socioeconomic status: a comparison of Spain and Denmark. *The Scandinavian Journal of Public Health*, 38(3 Suppl), 137-148. doi: 10.1177/1403494809353825
- Moncada, S., Utzet, M., Llorens, C., Galtés, A. y Moreno, N. (2010). *Encuesta de riesgos psicosociales en el trabajo en España, 2010. Trabajo de campo para la producción de la versión española del Copenhagen Psychosocial Questionnaire*. Madrid: Fundación de Prevención de Riesgos Laborales.

- Moncada, S., Utzet, M., Llorens, C., Galtés, A., Moreno, N. y Navarro, A. (2011). *Evolución de la exposición de los asalariados a los riesgos psicosociales en España entre los años 2005 y 2010*. Barcelona: ISTAS. Recuperado de http://www.copsoq.istas21.net/index.asp?ra_id=59
- Moncada, S., Utzet, M., Molinero, E., Llorens, C., Moreno, N., Galtés, A. y Navarro, A. (2014). The Copenhagen Psychosocial Questionnaire II (COPSOQ II) in Spain—A Tool for Psychosocial Risk Assessment at the Workplace. *American Journal of Industrial Medicine*, 57, 97–107 doi: 10.1002/ajim.22238
- Monk, T. y Folkard, S. (1985). Individual differences in shiftwork adjustment. En S. Folkard y T. Monk (Eds.), *Hours of work. Temporal factors in work scheduling* (pp. 227-237). Chichester: John Wiley & Sons.
- Moorman, R. H. y Podsakoff, P. (1992). A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. *Journal of Occupational and Organizational Psychology*, 65 (3), 131-149.
- Morales, P. (2006). *Medición de actitudes en psicología y educación. Construcción de escalas y problemas metodológicos*. (3ªed.). Madrid: Universidad Pontificia de Comillas.
- Morales, P. (2011). Guía para construir cuestionarios y escalas de actitudes. Madrid: Universidad Pontificia de Comillas. Recuperado de <http://web.upcomillas.es/personal/peter/otrosdocumentos/Guiaparaconstruirescalasdeactitudes.pdf>
- Morales, P. (2012a). *Análisis de ítems en las pruebas objetivas*. Universidad Pontificia Comillas, Facultad de ciencias humanas y sociales, Madrid, España. Recuperado de <http://web.upcomillas.es/personal/peter/otrosdocumentos/AnalisisItemsPruebasObjetivas.pdf>
- Morales, P. (2012b). *Estadística aplicada a las Ciencias Sociales*. Madrid: Biblioteca Univesidad Pontificia Comillas, Colección Ciencias Sociales, 03.
- Morales-Vives, F., Camps, E. y Lorenzo-Seva, U. (2013). Development and validation of Psychosocial Maturity Assessment Scale (PSYMAS). *European Journal of Psychological Assessment*, 29, 12-18.

- Morán, C. (2007). Acoso psicológico y burnout en profesionales de recursos humanos. En J. Romay (Ed.). *Perspectivas y retrospectivas de la Psicología Social en los albores del siglo XXI* (pp. 249-256). Madrid: Biblioteca Nueva.
- Moreno, B. (2004). Relaciones contractuales y acoso psicológico en el trabajo: un análisis comparativo Contractual. *Medicina Y Seguridad Del Trabajo, L* (197), 19–26.
- Moreno, B. (2011). Factores y riesgos laborales psicosociales: conceptualización, historia y cambios actuales. *Medicina y Seguridad en el Trabajo*, 57(1), 4–19. doi: 10.4321/S0465-546X2011000500002.
- Moreno, B. y Báez, C. (2009). *Análisis comparado de los métodos de evaluación de riesgos psicosociales. Balance de Perspectiva*. Observatorio Permanente de Riesgos Psicosociales de UGT. Universidad Autónoma de Madrid.
- Moreno, B. y Báez, C. (2010). *Factores y Riesgos Psicosociales, formas, consecuencias, medidas y buenas prácticas*. Instituto Nacional de Seguridad e Higiene en el trabajo. Madrid. Recuperado de <http://www.insht.es/InshtWeb/Contenidos/Documentacion/PUBLICACIONES%20PROFESIONALES/factores%20riesgos%20psico.pdf>
- Moreno, B., Bustos, R., Matallana, A. y Miralles, M. (1996). Validación del Cuestionario Breve de Burnout (CBB) en el ámbito educativo. *Comunicación presentada en el I Congreso de la Sociedad Española para el Estudio de la Ansiedad y el Estrés*. Benidorm.
- Moreno, B., Gálvez, M., Garrosa, E. y Mingote, J.C. (2006). Nuevos planteamientos en la evaluación del burnout. La evaluación específica del desgaste profesional médico. *Atención Primaria*, 38(10), 544-549.
- Moreno, B. y Garrosa, E. (2013). *Salud laboral*. Madrid: Pirámide.
- Moreno, B., Garrosa, E. y González, J.L. (2000). El desgaste profesional de enfermería: Desarrollo y validación factorial del CDPE. *Archivos de Prevención de Riesgos Laborales*, 3, 18-28.
- Moreno, B., Garrosa, E. y González, J.L. (2001). La evaluación del estrés y el burnout del profesorado: el CBP-R. *Revista de Psicología del Trabajo y de las Organizaciones*, 16(1), 131-349.

- Moreno, B., Oliver, C. y Aragonese, A. (1991). El burnout, una forma específica de estrés laboral. En Buela-Casal, G. y Caballo, V. (Eds). *Manual de psicología clínica aplicada* (pp. 271-285). Madrid. Siglo XXI.
- Moreno, B., Oliver, C. y Aragonese, A. (1993). *Configuración específica del estrés laboral asistencial en el profesorado de educación media*. Madrid, España: Centro Nacional de Investigación Educativa (CIDE).
- Moreno, B., Rodríguez-Muñoz, A., Morante, M. E., Garrosa, E., Rodríguez-Carvajal, R. y Díaz-Gracia, L. (2008). Evaluación del acoso psicológico en el trabajo: Desarrollo y estudio exploratorio de una escala de medida. *Universitas Psychologica*, 7, 335-345.
- Moreno, V., Traba, J. y Morales, M.B. (2008): *Las medidas agroambientales y la conservación de las aves esteparias. Análisis de eficiencia y propuesta de mejora para las Estepas Cerealistas de los ríos Jarama y Henares* (Madrid). Servicio de Publicaciones. Universidad Autónoma de Madrid
- Moreno-Rosset, C. (2005) (Edit.). *Evaluación psicológica. Concepto, proceso y aplicación en las áreas del desarrollo y de la inteligencia*. (2ª ed.). Madrid: Sanz y Torres.
- Morillas, R. M., Rubio-Romero, J. C. y Fuertes, A. (2013). A comparative analysis of occupational health and safety risk prevention practices in Sweden and Spain. *Journal of Safety Research*, 47, 57–65.
- Morris, C. G y A. A. Maisto (2005). *Introducción a la psicología* (12ª ed.). Mexico: Pearson Education.
- Morse, J. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, 40, 120-123
- Morse, D. T. y Morse, L. W. (2002). Are undergraduate examinee's perceptions of item difficulty related to item characteristics? *Perceptual and Motor Skills*, 95(3-2), 1281-1286.
- Moya, L. M. (2009). *Introducción a la estadística de la salud*. (2. ed). San José, C.R.: Editorial Universidad de Costa Rica.

- Mueller-Hanson, R., Heggstad, E. D. y Thornton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, 88, 348–355. doi:10.1037/0021-9010.88.2.348
- Mullen, P. M. (2003). Delphi: myths and reality. *Journal of Health Organisation and Management*, 17(1), 37-52.
- Muniamuthu, S. y Raju, R. (2010). The Influence of Psychosocial Factors on the Occurrence of Musculoskeletal Disorders Regarding VDT Users, *European Journal of Scientific Research*, 43(2), 291-293.
- Muñiz, J. (1998a). *La teoría clásica de los tests*. Madrid: Pirámide.
- Muñiz, J. (1998b). La medición de lo psicológico. *Psicothema*, 10(1), 1–21.
- Muñiz, J. (2004). La validación de los tests. *Metodología de las Ciencias del Comportamiento*, 5, 121-141.
- Muñiz, J. (2005). La validez desde una óptica psicométrica. *Acta Comportamentalia*, 13(1), 9-20.
- Muñiz, J. (2010). Las teorías de los test: Teoría clásica y teoría de respuesta de los ítems. *Papeles del Psicólogo*, 31(1), 57-66.
- Muñiz, J. y Bartram, D. (2007). Improving international tests and testing. *European Psychologist*, 12, 206-219.
- Muñiz, J., Fidalgo, A. M., García-Cueto, E., Martínez, R. y Moreno, R. (2005). *Análisis de los ítems*. Madrid: La Muralla.
- Murray, M., Fitzpatrick, D. y O'Connell, C. (1997). Fishermen's blues: Factors related to accidents and safety among Newfoundland fishermen. *Work & Stress*, 11(3), 292-297.
- Murry, J. y Hammons, J. (1995). Delphi: A versatile methodology for conducting qualitative research. *The Review of Higher Education*, 18(4), 423-436.
- Muthen, B. y Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology*, 45(1), 19-30.

- Narayan, S. y Krosnick, J. A. (1996). Education moderates some response effects in attitude measurement. *Public Opinion Quarterly*, 60, 58-88.
- Nardi, P. M. (2003). *Doing Survey Research: A Guide to Quantitative Methods*. Boston: Allyn and Bacon.
- Nater, U.M., Moor, C., Okere, U., Stallkamp, R., Martin, M., Ehlert, U. y Kliegel, M. (2007). Performance on a declarative memory task is better in high than low cortisol responders to psychosocial stress. *Psychoneuroendocrinology*, 32, 758-763.
- Navarro-Gonzalez, D., Vigil-Colet, A., Ferrando, P y Lorenzo-Seva, U. (2016). *Taller: Psychological Test Toolbox: Una nueva herramienta para calcular el análisis factorial controlando sesgos de respuesta en test autoadministrados*. Tarragona: Universitat Rovira i Virgili.
- Navas, M. J. (1999). Un siglo utilizando tests. *Revista Electrónica de Metodología Aplicada*, 4(2), 1-11.
- Navas, M. J. (Coord.) (2010). *Métodos, diseños y técnicas de investigación psicológica* (Ed. digital). Madrid: UNED.
- Nawas, M. (1971). El Estilo de Vida II. Estilos cognoscitivos. *Revista Latinoamericana de Psicología*, 3(2), 191-210.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15, 263-280.
- Neill, J. A. y Jackson, D.N. (1970). An evaluation of item selection strategies in personality scale construction. *Educational and Psychological Measurement*, 30, 647-661.
- Nelms, K. R. y Porter, A. L. (1991). EFTE: an interactive delphi method. *Technological, Foresicasting and Social Change*, 28, 43-61
- Netterstrom, B., Conrad, N., Bech, P., Fink, P., Olsen, O., Rugulies, R. y Stansfeld, S. (2008). The relation between work-related psychosocial factors and the development of depression. *Epidemiologic Reviews* 30, 118-132.

- Nick, J. y Lee-Ross, D. (1998). *Research methods in service industry management*. London: Cassell.
- Niedhammer, I., Goldberg, M., Leclerc, A., Bugel, I. y David, S. (1998). Psychosocial factors at work and subsequent depressive symptoms in the Gazel cohort. *Scand J Work Environ Health*, 24, 197–205.
- Niño, J. (2002). Evaluación de los riesgos laborales y factores psicosociales. *MAPFRE Seguridad* 85, 25-35.
- Niño, J. (2006). Los errores en las evaluaciones de los riesgos psicosociales: prevenir la confusión. *MAPFRE. Seguridad* 103, 28-41.
- Ng, D. M. y Jeffery, R. W. (2003). Relationships between perceived stress and health behaviors in a sample of working adults. *Health Psychology*, 22(6), 638–642.
- Noëlle-Neumann, E (1995). *La espiral del silencio. Opinión pública: nuestra piel social*, Paidós. Barcelona.
- Nogareda, S. (2001). *Estrés en el colectivo docente; metodología para su evaluación*. (NTP 574). Madrid: INSHT. Recuperado de :http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/501a600/ntp_574.pdf
- Nogareda, C., Gracia, E., Martínez, I. M. y Salanova, M. (2007). *El trabajo emocional: concepto y prevención*. NTP N° 720. Madrid: INSHT
- Nordin, M., Knutsson, A., Sundbom, E. y Stegmayr, B. (2005). Psychosocial factors, gender, and sleep. *Journal of Occupational Health Psychology*, 10(1), 54–63.
- Nuebling, M., Stöbe, U., Hasselhorn, H-M., Michaelis, M. y Hofmann, F. (2006) *Measuring psychological stress and strain at work: Evaluation of the COPSQ Questionnaire in Germany*. *GMS Psycho-Social-Medicine* 3. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2736502/#R19>
- Nuevo, R., Montorio, I., Márquez-González, M., Cabrera, I., Izal, M. y Pérez-Rojo, G. (2009). Diferencias asociadas a la edad en el efecto de la deseabilidad social en el autoinforme del estado emocional. *Revista Española de Geriatria y Gerontología*, 44(2), 85-89.

- Nunnally, J. C. y Bernstein, I. J. (1995). *Teoría Psicométrica* (Tercera ed.). México: McGraw-Hill Interamericana.
- Nworie, J. (2011). Using the Delphi Technique in Educational Technology Research. *TechTrend* 55(5), 24-30.
- O'Brien, G.E. (1986). *Psychology of work and unemployment*. Chichester, UK: Wiley.
- OCDE (2017). *Estudios económicos de la OCDE: España, 2017*. Ediciones OCDE. París. doi: 10.1787/9789264271920-es
- Ochoa, E. y Madoz, A. (2008). Consumo de alcohol y otras drogas en el medio laboral. *Medicina y Seguridad del Trabajo*, 54(213), 25-32.
- Ohly, S. y Fritz, C. (2010). Work characteristics, challenge appraisal, creativity, and proactive behaviour: A multi-level study. *Journal of Organisational Behaviour*, 31(4), 543-565.
- Okoli, C. y Pawlowski, S. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, 42,15-29. doi:10.1016/j.im.2003.11.002
- Olea, J., Abad, F. J. y Barrada, J. R. (2010). Tests informatizados y otros nuevos tipos de tests. *Papeles del Psicólogo*, 31(1), 94-107.
- O'Leary, A. (1990). Stress, emotion and human immune function. *Psychological Bulletin*, 103, 363-382.
- Olesen, K. G., Thoft, E., Hasle, P. y Kristensen, T.S. (2008). *Virksomhedens sociale capital*. Hvidbog. København: Det Nationale Forskningscenter for Arbejdsmiljø (NFA)[The social capital of the companies. A white book].
- Oliver, A., Galiana, L., Sancho, P. y Tomás, J. (2011). ¿Evitando la aquiescencia se produce sesgo de método?: un estudio a través de escalas de autoestima. *Anuario de psicología*, 41(1-3), 155-169.
- Oliver, J. E., Mansell, A. y Jose, P.E. (2010). A longitudinal study of the role of negative affectivity on the work stressor-strain process. *International Journal of Stress Management*, 17(1), 56-77.

- Olobatuyi, M. E. (2006). *A user's guide to path analysis*. Lanham, MD: University Press of America
- Oncins, M. y Almodóvar, A. (1997). *Factores Psicosociales: fases para su evaluación (NTP 450)*. Madrid: INSHT.
- Organización Internacional del Trabajo (O.I.T.) (1984). *Factores psicosociales en el trabajo: naturaleza, incidencia y prevención*. Ginebra. Informe del Comité Mixto OIT-OMS sobre medicina del trabajo, novena reunión Ginebra, 24 de septiembre de 1984.
- Organización Internacional del Trabajo (O.I.T.) (1986). *Factores psicosociales en el trabajo: reconocimiento y control*. Ginebra: OIT.
- Organización Internacional del Trabajo. (2000). S.O.S. Estrés en el trabajo: aumentan los costos del estrés en el trabajo y la incidencia de la depresión es cada vez mayor. *Revista de la OIT*, 37, 4-5.
- Organización Internacional del Trabajo (O.I.T.) (2012). *SOLVE: integrando la promoción de la salud a las políticas de SST en el lugar de trabajo: guía del formador*. Ginebra: OIT.
- Organización Internacional del Trabajo (O.I.T.) (2016). *Estrés en el trabajo: un reto colectivo*. Turín: Centro internacional de formación de la OIT. Recuperado de http://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---safework/documents/publication/wcms_466549.pdf
- Organización Mundial de la Salud, OMS. (1948). Constitución de la Organización Mundial de la Salud. En *Official records of the World Health Organization*, nº 2, p. 100. Ginebra: World Health Organization.
- Organización Mundial de la Salud, OMS. (2004). *La organización del trabajo y el estrés*. Serie protección de la salud de los trabajadores Nº 3. Francia.
- Organización Mundial de la Salud, OMS. (2010). *Entornos laborales saludables fundamentos y modelo de la OMS contextualización, prácticas y literatura de apoyo*. OMS: Suiza. Recuperado de http://www.who.int/occupational_health/evelyn_hwp_spanish.pdf
- Ortega, F. (2008) El método Delphi, prospectiva en ciencias Sociales a través del análisis de un caso práctico. *Revista Escuela de Administración de Negocios*, 65, 31-64.

- Óscar, A., González-Camino, G., Bardera, P. y Peiró, J.M. (2003). Estrés de rol y su influencia sobre el bienestar psíquico y físico en soldados profesionales. *Psicothema*, 15(1), 54-57.
- Osipow, S. H. (1998). *A manual for the Occupational Stress Inventory Revised Edition (OSI-R)*. (Professional Manual). Lutz, Florida: Psychological Assessment Resources.
- Ostry, A., Maggi, S., Tansey, J., Dunn, J., Hershler, R., Chen, L...Hertzman, C. (2007). The impact of psychosocial work conditions on attempted and completed suicide among western Canadian sawmill workers. *The Scandinavian Journal of Public Health*, 35, 265–271.
- Paar, G. H., Bezenberger, U. y Lorenz-Meyer, H. (1988). The correlation of psychosocial stress and disease activity in patients with Crohn disease and ulcerative colitis. *Z. Gastroenterol*, 26(10), 648–657.
- Pace, V. L. y Borman, W. C. (2006). The use of warning to discourage faking on noncognitive inventories. En R. L. Griffith y M. H. Peterson (Eds.). *A closer examination of applicant faking behavior* (pp. 283-304). Greenwich, Connecticut: Information Age Publishing.
- Padilla, J. L., Gómez, J., Hidalgo, M. D. y Muñoz, J. (2006). La evaluación de las consecuencias del uso de los tests en la teoría de la validez. *Psicothema*, 19, 307-312.
- Padrós, F., Soria-Mas, C. y Navarro, G. (2012). Afecto positivo y negativo ¿Una dimensión bipolar o dos dimensiones unipolares independientes?; Positive and negative affect: One bipolar dimension or two independent unipolar dimensions? *Interdisciplinaria*, 29 (1), 151- 164. Recuperado de http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1668-70272012000100009
- Pando, M., Carrión, M. A., Arellano, G. y Saraz, S. (2006.) *Los factores psicosociales en el trabajo en Factores Psicosociales y Salud Mental en el Trabajo*. Universidad de Guadalajara. México.
- Paraskevas, A. y Saunders, M. N. K. (2012). Beyond consensus: an alternative use of Delphi enquiry in hospitality research. *International Journal of Contemporary Hospitality Management*, 24(6), 907–924. doi: 10.1108/09596111211247236

- Park, B. (1998) Factors influencing psoriasis: An analysis based upon the extent of involvement and clinical type. *The Journal of Dermatology*, 25(2), 97-102.
- Parker, D. F. y De Cotiis T. A. (1983). Organizational Determinants of Job Stress. *Organizational Behavior and Human Performance*, 32, 160-177.
- Paulhus, D. L. (1981). Control of social desirability in personality inventories: Principal-factor deletion. *Journal of Research in Personality*, 15, 383-388.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. En A. Angleitner y J. S. Wiggins (Eds.), *Personality Assessment Via Questionnaires* (pp. 143-165). New York: Springer-Verlag.
- Paulhus, D. L. (1991). Measurement and control of response bias. En J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. En H. I. Braun y D.N. Jackson (Eds.), *Role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah, NJ: Lawrence Erlbaum Associates.
- Paulhus, D. L. y John, O. (1998). The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, 66, 1025-1060.
- Paulhus, D. y Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, 60, 307– 317.
- Paulhus, D. L. y Vazire, S. (2007). The self-report method. En R. W. Robins, R. C. Fraley y R.F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224–239). New York: Guilford 7.
- Pauls, C. A. y Stemmler, G. (2003). Substance and bias in social desirability responding. *Personality and Individual Differences*, 35, 263-275.

- Pavlou, P. A., Liang, H. y Xue, Y. (2007). Understanding and mitigating uncertainty in online exchange relationships: A principal-agent perspective. *MIS Quarterly*, 31 (1), 105-136.
- Payne, N., Jones, F. y Harris, P. (2002). The impact of working life on health behaviour: the effect of job strain on the cognitive predictors of exercise. *Journal of Occupational Health Psychology*, 7(4), 342-353.
- Peer, E. y Gamliel, E. (2011). Too reliable to be true? Response bias as a potential source of inflation in paper-and-pencil questionnaire reliability. *A Peer-Reviewed Electronic Journal*, 16(9), 1-8.
- Peiró, J. M. (1999a). *Desencadenantes del estrés Laboral*. Madrid: Pirámide.
- Peiró, J. M. (1999b). El Modelo AMIGO: marco contextualizador del desarrollo y la gestión de recursos humanos en las organizaciones. *Papeles del Psicólogo*, 72, 3-15.
- Peiró, J. M. (2000). *Desencadenantes del estrés laboral*. Madrid: Pirámide.
- Peiró, J. M. (2001). El estrés laboral: una perspectiva laboral y colectiva. *Prevención, Trabajo y Salud*, 13, 1-12.
- Peiró, J. M. (2004). El Sistema de Trabajo y sus implicaciones para la prevención de los riesgos psicosociales en el trabajo. *Universitas Psychologica*, 3(2), 179-186.
- Peiró, J.M. (2005). *Desencadenantes del estrés laboral*. Madrid: Pirámide.
- Peiró, J. M. (2006). Metodología Prevenlab-psicosocial: Universidad de Valencia. En J. L. Meliá, C. Nogareda, M. Lahera, A. Duro, J. M. Peiró, R. Pou y otros, *Perspectivas de intervención en Riesgos Psicosociales: Evaluación de riesgos* (pp. 105-130). Foment del Treball Nacional.
- Peiró, J. M. (2008). *El Modelo Amigo: Un "Mapa" Conceptual para evaluar los Riesgos Psicosociales y gestionar su Prevención*. Universidad de Valencia: FOCAD.
- Peiró, J. M. (2009). Nuevas tendencias en la investigación sobre el estrés laboral y sus implicaciones para el análisis y prevención de los riesgos sociales. *Lección Magistral leída en solemne acto del curso 2009-2010*. Universidad de Valencia.

- Peiró, J. M. (2010). Cuestiones fundamentales en la evaluación de los riesgos psicosociales: Avances y resistencias para su clarificación. *Prevención De Riesgos Psicosociales y Bienestar en el Trabajo*, 2, 13-66.
- Peiró, J.M. y Salvador, A. (1993). *Control del estrés laboral*. Madrid. Eudema.
- Peiró, J.M., J. Yeves y L. Lorente (2013). El análisis de los riesgos psicosociales en el trabajo: Investigación y práctica profesional en España. En Rufino, M. (dir.): *Anuario internacional sobre prevención de riesgos psicosociales y calidad de vida en el trabajo*. Madrid: Secretaría de Salud Laboral y Medio Ambiente UGT-CEC, pp. 19-46.
- Pejtersen, J. H, Kristensen, T. S., Borg, V. y Bjorner, J. B. (2010). The second version of the Copenhagen Psychosocial Questionnaire (COPSOQ II). *Scandinavian Journal Public Health*, 38(3), 8-24.
- Peled, R., Carmil, D., Siboni-Samocha, O. y Shoham-Vardi, I. (2008). Breast cancer, psychological distress and life events among young women. *BMC Cancer*, 8, 245. doi: 10.1186/1471-2407-8-245
- Pennebaker, J. W. (1989). Confession, inhibition, and disease. En L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 211-244). San Diego, CA: Academic Press.
- Pennebaker, J. W., Rime, B. y Blankenship (1996). Stereotypes of emotional expressiveness of Northerners and Southerners. *Journal of Personality and Social Psychology*, 70, 372-380.
- Peña, A., Ramos, M. J. y Martínez-Losa, J. F. (2009). Valoración psicosocial mediante el uso de los modelos de estrés laboral de y desequilibrio. En A. Mondelo, P; Mattila, M.; Karwowski, W.; Hale (Ed.). *Proceedings of the Sixth International Conference on Occupational Risk Prevention - ORP2008*, 51, 1 - 11. Recuperado de http://fesprosa.org.ar/portal/wp-content/uploads/2014/08/Pena_Pulido_y_colaboradores1.pdf
- Pérez-Amoros, F. (2015). Trabajador autónomo económicamente dependiente en el ordenamiento español una figura controvertida y contradictoria. *Derecho laboral: Revista de doctrina, jurisprudencia e informaciones sociales*, 257, 5-30.

- Pérez-Andrés C. (2000) ¿Deben estar las técnicas de consenso incluidas entre las técnicas de investigación cualitativa? *Revista Española de Salud Pública*, 74, 319-321.
- Pérez, C., Rodríguez, M. J., Guerrero, A., Margarit, C., Martín-Estefanía, C., Oteo-Álvaro, A. y Caballero, F. (2013). Consenso experto sobre el uso clínico de los tratamientos por vía tópica en el manejo del dolor neuropático periférico. *Revista de La Sociedad Española del Dolor*, 20(6), 308-323. doi: 10.4321/S1134-80462013000600005
- Pérez, I., Torres, E., Alcorta, I., Etxeberria, A., Rotaecche del Campo, R. y Reviriego, E. (2009). Exploración de barreras y facilitadores para la implementación de guías de práctica clínica: un estudio Delphi. Plan de Calidad para el Sistema Nacional de Salud del Ministerio de Sanidad y Política Social. Servicio de Evaluación de Tecnologías Sanitarias del País Vasco (Osteba.; *Informes de Evaluación de Tecnologías Sanitarias: OSTEBA N° 2007/0*.
- Peró, M. y Guàrdia, J. (2010). *Algunas aproximaciones a las técnicas de evaluación de riesgos psicosociales*. En J. Guàrdia (Coord). *La evaluación de riesgos psicosociales. Guía de buenas prácticas. Métodos de evaluación y sistemas de gestión de riesgos psicosociales: un balance de utilidades y limitaciones* (pp. 41-52). Jaén: Blanca Impresiones
- Peró, M. (2010). Análisis de datos. En J. Guàrdia (Coord). *La evaluación de riesgos psicosociales. Guía de buenas prácticas. Métodos de evaluación y sistemas de gestión de riesgos psicosociales: un balance de utilidades y limitaciones* (p. 93). Jaén: Blanca Impresiones
- Perry-Jenkins, M., Goldberg, A., Pierce, C. y Sayer, A. (2007). Shift work, role overload, and the transition to parenthood. *Journal of Marriage and Family*, 69(2), 123–138.
- Peters, M. L., Godaert, G. L.R., Ballieux, R. E., Brosschot, J.F., Sweep, F.C.G.J., Swinkels, L. M... Heijnen, C.J. (1999). Immune responses to experimental stress: effects of mental effort and uncontrollability. *Psychosomatic Medicine* 61(4), 513-524.
- Pettit, F. A. (2002). A comparison of World-wide-web and paper-and-pencil personality questionnaires. *Behavior Research Methods*, 34(1), 50-54.
- Pikhart, H., Bobak, M., Pajak, A., Malyutina, S., Kubinova, R., Topór-Madry, R., ... Marmot, M. (2004). Psychosocial factors at work and depression in three countries of Central and Eastern Europe. *Social Science & Medicine*, 58, 1475-1482.

- Pikhart, H., Bobak, M., Siegrist, J., Pajak, A., Rywik, S., Khyshegye, J., ... y Marmot, M. (2001). Psychosocial work characteristics and self-rated health in four postcommunist countries. *Journal of Epidemiology and Community Health*, 55, 624-630.
- Pill, J. (1971). The Delphi method: substance, context, a critique and an annotated bibliography. *Socio-Economic Planning and Science*, 5, 57-71.
- Pines, A. M. y Aronson, E. (1988). *Carreer burn out: causes and cures*. Nueva York (Free Press).
- Pines, A. M., Aronson, E. y Kafry, D. (1981). *Burnout: From tedium to personal growth*. New York: Free Press.
- Pingitore, G., Chrobak, V. y Petrie, J. (1991). The social and psychologic factors of bruxism. *The Journal of Prosthetic Dentistry*, 65(3), 443-446.
- Piñuel, I. (2001). *Mobbing: cómo sobrevivir al acoso psicológico en el trabajo*. Santander: Sal Terrae.
- Piñuel, I. (2002). *La incidencia del mobbing o acoso psicológico en el trabajo en España. Informe Cisneros II sobre violencia en el entorno laboral*. Madrid: Universidad de Alcalá de Henares.
- Piñuel, I. (2008). El mobbing. La evaluación y la prevención del acoso como riesgo psicosocial en el trabajo, En M. Clemente (Dir.), *Manual de Psicología jurídica laboral* (pp. 51-85) Madrid: Delta Publicaciones.
- Piñuel, I. y Oñate, A. (2006). La evaluación y diagnóstico del mobbing o acoso psicológico en la organización: el barómetro Cisneros. *Revista de Psicología Del Trabajo y de Las Organizaciones*, 22(3), 309-332.
- Plans, B. y León, O. (2003) ¿Como debe ser el doctorando ideal en Psicología? Contesta al oráculo de Delfos. *Psicothema*, 15(4), 610-614.
- Pocinho, M. y Perestrelo, C. X. (2011). Um ensaio sobre burnout, engagement e estratégias de coping na profissão docente. *Educação e Pesquisa*, 37(3), 513-528.

- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y. y Podsakoff, N. P. (2003). *Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies*. *Journal of Applied Psychology*, 88(5), 879-903. doi: 10.1037/0021-9010.88.5.879
- Podsakoff, P.M., MacKenzie, S.B. y Podsakoff, N.P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 65, 539-569.
- Podsakoff, P.M. y Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12, 531–544. doi: 10.1177/014920638601200408
- Polaino-Lorente, A. y Martínez Cano, P. (2003). *Evaluación psicológica y psicopatológica de la familia*. Rialp. 2ª ed. Madrid.
- Ponsoda, V. (2009). Metodología al servicio del psicólogo. *Papeles del Psicólogo*, 31(1), 2-6 <http://www.papelesdelpsicologo.es/pdf/1791.pdf>
- Poo, F. (2007). El sesgo de deseabilidad social, de estilo de respuesta a rasgo de personalidad. *Anuario de Proyectos e Informes de Becarios de Investigación*, año 2007, 58-65. Recuperado de <http://www.mdp.edu.ar/psicologia/psico/investigacion/filesinves/anuario2007.pdf>
- Porta, M. (Ed.) (2014). *A dictionary of epidemiology* (6th edn.) New York: Oxford University Press.
- Porthé, V., Benavides, F., Vázquez, M. L., Ruiz-Frutos, C., García, A. M, Ahonen, E., ... Benach, J. (2009). La precariedad laboral en inmigrantes en situación irregular en España y su relación con la salud. *Gaceta Sanitaria*, 23(1), 107–114.
- Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing*, 41(4), 376-382, doi: 10.1046/j.1365-2648.2003.02537.x
- Prat, R. y Doval, E. (2005). Construcción y análisis estadístico de escalas. En J. Lèvy y J. Valera (Comp.) *Análisis Multivariante para las Ciencias Sociales*. Madrid: Pearson Prentice Hal.
- Press, J. E. y Townsley, E. (1998). Wives' and husbands' housework reporting: gender, class, and social desirability. *Gender & Society*, 12, 188–218.

- Prewett, M. S., Tett, R. y Christiansen, N. (2013). A Review and Comparison of 12 Personality Inventories on Key Psychometric Characteristics. En N. Christiansen & Rob Tett (Eds.), *The Handbook of Personality at Work*. Danvers, MA: Routledge.
- Prieto, G. y Delgado, A.R. (2010). Fiabilidad y Validez. *Papeles del Psicólogo*, 31(1), 67-74. Recuperado de <http://www.papelesdelpsicologo.es/pdf/1797.pdf>
- Prospéro, O., Méndez, M., Ruiz, A.E., Alvarado, I. y León, L. (2011). Insomnio, estrés y cannabinoides. *Salud Mental*, 34(3), 211 - 218.
- Putnam, R. D. (1993). The Prosperous Community: Social Capital and Public Life. *American Prospect*, 13, 35-42.
- Quade, E. S. (1967). *Cost Effectiveness: Some Trends In Analysis*. Santa Mónica, C.A.: The RAND Corporation.
- Quiceno, J. M. y Vinaccia, S. (2007). Redalyc.Burnout: “síndrome de quemarse en el trabajo (SQT). *Acta Colombiana de Psicología*, 10(2), 117–125.
- Quick, J. C. (Ed.). (1997). *Preventive stress management in organizations*. Washington: American Psychological Association.
- Quinlan, M. y Mayhew, C. B. P. (2001). The global expansion of precarious employment, work disorganization, and consequences for occupational health: a review of recent research. *International Journal of Health Services*, 31, 335-414.
- Rabada, I. y Artazcoz, L. (2002). Identificación de los factores de riesgo laboral en docentes: un estudio delphi. *Archivos de Prevención de Riesgos Laborales*, 5, 53-61.
- Räihä, I., Kemppainen, H., Kaprio, J., Koskenvuo, M. y Sourander, L (1988). Lifestyle, stress, and genes in peptic ulcer disease: a nationwide twin cohort study. *Archives of Internal Medicine*, 158(7), 698-704.
- Rao, S. K., Bhat, M. y David, J. (2011). *Work, stress, and diurnal bruxism: A pilot study among information technology professionals in Bangalore City*. India: Hindawi Publishing Corporation. doi:10.1155/2011/65489

- Ray, J. J. (1983). Reviving the problem of acquiescent response bias. *The Journal of Social Psychology*, 121(1), 81-96.
- Ray, J. J. (1990). *Acquiescent Response Tendency: An update and some data on the invalidity of the Dogmatism scale*. No publicado. Recuperado <http://jonjayray.tripod.com/acq.html>
- Razavi T. (2001). Self-report measures: An overview of concerns and limitations of questionnaire use in occupation stress research. *Discussion Papers in Accounting and Management Science*. Southampton, UK University of Southampton 23pp. Recuperado de <http://eprints.soton.ac.uk/35712>.
- Real Decreto 39/1997, de 17 de enero, por el que se aprueba el Reglamento de los Servicios de Prevención. *Boletín Oficial del Estado*. Madrid, 31 de enero de 1997, núm. 27, pp. 3031-3045.
- Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado. *Boletín Oficial del Estado*. Madrid, 10 de febrero de 2011, núm. 35, pp.13909-13926.
- Reavley, N. J., Ross, A., Killackey, E. J. y Jorm, A. F. (2012) Development of Guidelines to Assist Organisations to Support Employees Returning to Work after an Episode of Anxiety, Depression or a Related Disorder: A Delphi Consensus Study with Australian Professionals and Consumers. *BMC Psychiatry*, 3, 12-135.
- Reavley, N. J., Ross, A., Martin, A., LaMontagne, A. D. y Jorm, A. F. (2014). Development of guidelines for workplace prevention of mental health problems: A Delphi consensus study with Australian professionals and employees. *Mental Health and Prevention*, 2(1-2), 26–34. doi:10.1016/j.mhp.2014.07.002
- Reber, S. O. (2012). Stress and animal models of inflammatory bowel disease—an update on the role of the hypothalamo-pituitary-adrenal axis. *Psychoneuroendocrinology* 37(1),1–19 doi: 10.1016/j.psyneuen.2011.05.014
- Reguant-Álvarez, M. y Torrado-Fonseca, M. (2016). El método Delphi. *REIRE, Revista d'Innovació i Recerca en Educació*, 9(1), 87-102. doi:10.1344/reire2016.9.1916

- Regus Research Institute (2009). *Stress out? A study of trends in workplace stress across the globe*. Recuperado de http://www.regus.cn/images/Stress%20full%20report_FINAL_De_signed_tcm307-21560.pdf
- Regus (2012). *Workplace stress report 2012*. Recuperado de <http://press.regus.com/united-states/from-distressed-to-de-stressed/>
- Reidl, M. L., Guillén, R. R., Sierra, O. G. y Joya L. L. (2002). *Celos y envidia: medición alternativa*. México: Universidad Nacional Autónoma de México.
- Reig, A. y Caruana, A. (1990). Cuestionario de estrés ocupacional para profesionales de la salud: propiedades psicométricas preliminares. En J. M. Peiró (Coord.), Comunicaciones Area 3: Psicología del Trabajo y de las Organizaciones. *II Congreso del Colegio Oficial de Psicólogos*. Madrid: Colegio Oficial de Psicólogos.
- Restrepo, M. M. y Gómez-Restrepo, C. (2004). Metodología. *Revista Colombiana de Psiquiatría*, XXXIII(3), 327–335.
- Riaño, C. E. y Palomino, M. (2015). Diseño y elaboración de un cuestionario acorde con el método Delphi para seleccionar laboratorios virtuales (LV). *Sophia*, 11(2) 129-141.
- Rick, J., Briner, R. B., Daniels, K., Perryman, S. y Guppy, A. (2001). *A critical review of psychosocial hazard measures (HSE CRR 356)*. London: HSE Books.
- Rieger, W. (1986). Directions in Delphi Developments: Dissertations and Their Quality. *Technological Forecasting and Social Change*, 29,195-204.
- Risko, E. F., Quilty, L.C. y Oakman, J. M. (2006). Socially desirable responding on the web: Investigating the candor hypothesis. *Journal of Personality Assessment*, 87, 269-276.
- Rivas, F. (1979). *Elementos de Psicometría. Vol. I, II y III*. Valencia: Universidad de Valencia.
- Rivas, M. (2008). Salud y género: perspectiva de género en la salud laboral. *Revista Del Ministerio de Trabajo e Inmigración*, 74, 227–286.
- Robins, R. W., Fraley, R. C. y Kruger, R. F. (Eds.). (2007). *Handbook of Research Methods in Personality Psychology*. New York. Guilford Press.

- Robins, R. W., Hendin, H. M. y Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and social psychology bulletin*, 27(2), 151-161.
- Robinson, S. L. y Rousseau, D. M. (1994). Violating the Psychological Contract: Not the Exception but the Norm. *Journal of Organizational Behavior*, 15, 245 – 259.
- Robles, R. y Páez, F. (2003). Estudio sobre la traducción al español y las propiedades psicométricas de las escalas de afecto positivo y negativo (PANAS). *Salud mental*, 26(1), 69-75.
- Robson, S. M., Jones, A. y Abraham, J. (2008). Personality, Faking and Convergent Validity: A Warning Concerning Warning Statements. *Human Performance*, 21, 89-106.
- Rodríguez- Abuín, M. J. y González de Rivera, J. L. (2006). Diagnóstico diferencial del paranoïdismo en los síndromes de acoso. *Interpsiquis*, 1, 1–13.
- Rodríguez, D. e Ibiert, M. (2010). El método Delphi como método de investigación en la gestión de riesgos contra atentados terroristas. *XIV International Congress on Project Engineering*. Madrid 2010 Departamento Ingeniería Civil: Construcción, E.T.S.I.C.C.P. Universidad Politécnica de Madrid.
- Rodríguez, A. C., Martínez, N. A., Juárez, F., López, E. K., Carreño, S. y Medina-Mora, M. E. (2008). Relación entre el consumo de tabaco, salud mental y malestares físicos en hombres trabajadores de una empresa textil mexicana. *Salud Mental*, 31(4), 291-297.
- Romero, R., Cabero, J., Llorente M. C. y Vázquez-Martínez, A. I. (2012). El método Delphi y la formación del profesorado en TIC. *Global*, 9(44), 81-92.
- Rorer, L. G. (1965). The great response-style myth. *Psychological Bulletin*, 63(3), 129-156.
- Rorer, L. G. y Goldberg, L. R. (1965). Acquiescence in the MMPI? *Educational and Psychological Measurement*, 25(3), 801-817.
- Rosales, Y. y Rosales, F. R. (2013). Burnout estudiantil universitario. Conceptualización y estudio. *Salud Mental*, 36(4), 337–345.

- Rosas, A., Sánchez, J. y Chávez, M. M. (2012). La técnica Delphi y el análisis de la capacidad institucional de gobiernos locales que atienden el cambio climático. *Política y cultura*, 38, 165-194.
- Rospenda, K. M. y Richman, J. A. (2004). The factor structure of generalized workplace harassment. *Violence and Victims*, 19, 221-238.
- Rothman, K. J. (2012). *Epidemiology: An introduction*. (2nd ed.). New York: Oxford University Press.
- Rowe, G. y Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*, 15(4), 353-375.
- Rowe, G. y Wright, G. (1996). The impact of task characteristics on the performance of structured group forecasting technique. *International Journal of Forecasting*, 12(1), 73-89.
- Rowe, G. y Wright, G. (2001): Expert Opinions in Forecasting. Role of the Delphi Technique. En Armstrong (Ed.), *Principles of Forecasting: A Handbook of Researchers and Practitioners*. Boston: Kluwer Academic Publishers.
- Rowe, G., Wright, G. y Bolger, F. (1991). Delphi: A Reevaluation of Research and Theory. *Technological Forecasting and Social Change*, 39(3), 235-251.
- Rowe, G., Wright, G. y Mccoll, A. (2005). Judgment change during Delphi-like procedures: the role of majority influence, expertise, and confidence. *Technological Forecasting and Social Change*, 72, 377-399.
- Ruiz, E. e Idoate, V. (2005). *MPF cuestionario de factores psicosociales*. (Mini Psychosocial Factors). Pamplona, España: Autor.
- Ruiz, J. y Ispizua, M. A. (1989). La técnica Delphi. En J. Ruiz Olabuénaga, e M. A. Ispizua, La decodificación de la vida cotidiana. *Métodos de investigación cualitativa*. Bilbao, 171-179.
- Ruiz, M. A., Pardo, A. y San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del Psicólogo*, 31(1), 34 - 45.

- Ruiz-Pamies, M. (2014). I - DAQ: Desarrollo de un cuestionario libre de sesgos de respuesta para la evaluación de la agresividad. *Tesis Doctoral. Universidad Rovira i Virgili*.
- Ruiz-Pamies, M., Lorenzo-Seva, U., Morales-Vives, F. y Vigil-Colet, A. (2014). I-DAQ: A new test to assess direct and indirect aggression free of response bias. *The Spanish Journal of Psychology*, 17, 1-8. doi.:10.1017/sjp.2014.43
- Rundmo, T. (1994). Associations between safety and contingency measures and occupational accidents on offshore petroleum platforms. *Scandinavian Journal of Work, Environment & Health*, 20(2), 128-131.
- Rusting, C. L. y Larsen, R. J. (1998). Personality and cognitive processing of affective information. *Personality and Social Psychology Bulletin*, 24, 200-213.
- Sackeim, H. A. y Gur, R. C. (1978). Self-deception, selfconfrontation and consciousness. En G. E. Schwartz y D. Shapiro (Eds.), *Consciousness and self-regulation: advances in research and theory* (Vol. 2, pp. 139-197). New York, NY: John Wiley & Sons.
- Sackett, D. L. (1979.) Bias in analytic research. *J Chron Dis*, 32, 51-63.
- Saéz, M. C. y García-Izquierdo, M. (2009). Violencia psicológica en el trabajo: el mobbing. En J. Buendía y F. Ramos (Coord.), *Empleo, estrés y salud*. (pp. 191-204). Madrid: Pirámide.
- Salamanca, A. B. (2013). *El "AEIOU" de la investigación en Enfermería*. Madrid: FUDEN.
- Salanova, M., Bresó, E. y Schaufeli, W. (2005). Hacia un modelo espiral de las creencias de eficacia en el estudio de burnout y del engagement. *Revista Ansiedad y Estrés*, 11(2-3), 215-231.
- Salanova, M., Grau, R. M. y Martínez, I. M. (2005). Job demands and coping behaviour: the moderating role of professional self- efficacy. *Psicothema*, 17(3), 390-395.
- Salanova, M., Llorens, S., Cifre, E. y Martínez, I. (2006). Metodología RED-WoNT. En Foment del Treball Nacional (Ed.). *Perspectives de Intervenció en Riesgos Psicosociales, Evaluació de riscos*. (pp. 131-154). Barcelona.
- Salanova, M. y Schaufeli, W. (2009). *El engagement en el trabajo. Cuando el trabajo se convierte en pasión*. Madrid: Alianza.

- Salanova, M., Schaufeli, W. B., Llorens, S., Peiró, J. M. y Grau, R. (2000). Desde el “Burnout” al “Engagement”: ¿una nueva perspectiva? *Revista de Psicología del Trabajo y las Organizaciones*, 16(2), 117-134.
- Salas, C. y Melià, J. L. (2004). Evaluación de Riesgos Psicosociales y Factores Psicosociales que afectan a la Probabilidad de Accidente en una Empresa de Transporte Urbano de Economía Social. *Trabajo presentado al Tercer Congreso Internacional de Riesgos Laborales*. Santiago de Compostela.
- Saldarriaga, L. H. (2013). El contrato psicologico en las organizaciones, un fenomeno real. *Revista Poiésis*, 26, 1 – 10. Recuperado de <http://www.funlam.edu.co/revistas/index.php/poiesis/article/view/1018/906>
- Salgado, J. F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones. *Papeles del Psicólogo*, 26, 115-128.
- Salgado, J. F., Remeseiro, C. y Iglesias, M. (1996). Personality and Test-Taking Motivation. *Psicothema*, 8, 553-562.
- Sánchez, C. y Conde, P. (2008). La protección social y los riesgos psicosociales. *Anales del Derecho*, 26, 275–297.
- Sánchez, L., Simoes, M. D. y Brito, J. C. (2008). Trabajo y salud mental. Caso supervisores de una locación petrolera. *Salud de los trabajadores*, 16(1), 39–52.
- Sánchez-Díaz, S. (2013). Capacidad de trabajar bajo presión (resistencia al estrés y autocontrol). En *Como ser competente. Competencias profesionales demandadas en el mercado laboral* (pp. 82-88). Salamanca: Cátedra de Inserción Profesional Caja Rural de Salamanca. Univesidad de Salamanca.
- Sánchez-Urán, Y. (2004). Igualdad de género y salud laboral: propuestas institucionales, realidad normativa y práctica convencional. *Revista del Ministerio de Trabajo y Asuntos Sociales*, 53, 149-179.
- Sandi, C. y Calés, J. M. (2000). *Estrés: Consecuencias psicológicas, fisiológicas y clínicas*. Madrid: Editorial Sanz y Torres.

- Sandín, B. (1995). El estrés. En A. Belloch, B. Sandín y F. Ramos (eds.), *Manual de psicopatología*. (vol. 2, pp. 3-52). Madrid: McGraw-Hill.
- Sandín, B., Chorot, P., Lostao, L., Joiner, T., Santed, M. y Valiente, R. (1999). Escala PANAS de afecto positivo y negativo: validación factorial y convergencia transcultural. *Psicothema*, 11(1), 37-51.
- Santed, M. A., Sandín, B., Chorot, P., Olmedo, M. y Campayo, J. (2001). El papel de la afectividad negativa y positiva en las relaciones entre el estrés diario y la sintomatología somática: un estudio intra e interindividual. *Cuadernos de Medicina Psicosomática y Psiquiatría de Enlace*, 58/59, 41-53.
- Sanz, J., Izquierdo, A. y García-Vera, M. P. (2013). Una revisión desde la perspectiva de la validez de contenido de los cuestionarios, escalas e inventarios autoaplicados más utilizados en España para evaluar la depresión clínica en adultos. *Psicopatología Clínica, Legal y Forense*, 13, 139-175.
- Sapolsky, R. (2004). *Why Zebras Don't Get Ulcers: The Acclaimed Guide to Stress, Stress-Related Diseases, and Coping*. (3rd ed.) New York: Henry Holt & Co.
- Sassaroli, S. y Ruggiero, G. M. (2005). The role of stress in the association between low self esteem, perfectionism and worry and eating disorders. *International Journal of Eating Disorders*, 37(2), 135-141.
- Sauter, S. L., Hurrell, J. J. y Cooper, C. L. (1989). *Job Control and Worker Health*. Chichester, UK: Wiley and Sons.
- Sauter, S. L., Hurrell Jr., J. J., Murphy, L. R. y Leví, L. (Dirs.) (2001). *Factores Psicosociales y de Organización*. Capítulo 34 de la Enciclopedia de Salud y Seguridad en el Trabajo. Vol. II. Recuperado de <http://www.insht.es/InshtWeb/Contenidos/Documentacion/TextosOnline/EnciclopediaOIT/tomo2/34.pdf>
- Sauter, S. L., Murphy, L. R. y Hurrell, J. J. (1992) Prevention of work related psychological disorders: a national strategy proposed by the National Institute for Occupational Safety and Health. En G. P. Keita y S. L. Sauter (eds.) *Work and Well-Being: An Agenda for the 1990s*. American Psychological Association, Washington DC.

- Schaffin, W. W. y Talley, W. K., (1980). Individual stability in Delphi studies. *Technological Forecasting and Social Change*, 16, 70-73.
- Schaufeli, W. B. y Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study. *Journal of Organizational Behavior*, 25, 293-315.
- Schaufeli, W. B., Leiter, M. P, Maslach, C. y Jackson, S. E. (1996). Maslach Burnout Inventory - General survey. En C. Maslach, S E. Jackson & M.P. Leiter (Eds.), *The Maslach Burnout Inventory-Test Manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press
- Schaufeli, W.B. y Salanova, M. (2002). La evaluación de riesgos psicosociales en el trabajo. *Revista de Prevención, Trabajo y Salud*, 20, 4-9.
- Scheibe, M., Skutsch, M. y Schofer, J., (1975). *Experiment in Delphi Methodology*, en: *The Delphi Method: Techniques and Applications*. Massachussets.
- Schein, E. (1982). *Psicología de la Organización*. México: Editorial Prentice Hall Hispanoamericana.
- Schleyer, T. K. y Forrest, J. L. (2000). Methods for the design and administration of web-based surveys. *Journal American Medical Informatics Association*, 7, 416-425.
- Schnall, P. L., Belkic, K., Landsbergis, P. y Baker, D. (Eds.). (2000). Why the workplace and cardiovascular disease? *Occupational Medicine: State of the Art Reviews*, 15, 1-334.
- Schnall, P. L., Landsbergis, P. A. y Baker, D. (1994). Job strain and cardio-vascular disease. *Annual Review of Public Health*, 15, 381-411.
- Schoenbach, V. J. y Rosamond, W. D. (2000). *Fundamentals of epidemiology: An evolving text*. Chapel Hill: University of North Carolina. Schoenwald, S. K.
- Schriesheim, C. A. y Hill, K. D. (1981). Controlling Acquiescence Response Bias by Item Reversals: The Effect on Questionnaire Validity. *Educational and Psychological Measurement*, 41(4), 1101-1114. doi:10.1177/001316448104100420
- Schuessler, K, Hittle, D. y Cardascia, J. (1978). Measuring responding desirably with attitudeopinion items. *Social Psychology*, 41, 224-235.

- Schulz, W. y Brese, F. (2008). Assessing Student Knowledge, Background and Perception in the International Civic and Citizenship Education Study, Publicación presentada en *Annual Meeting of the American Educational Research Association (AERA)* in New York, 24-28 March.
- Schuman, H. y Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.
- Segerstrom, S. C. y Miller, G.E. (2004). Psychological stress and the human immune system: A meta-analytic study of 30 years of inquiry. *Psychological Bulletin*, 130, 601–641.
- Seguí, E. y Server, R. J. (2010). El capital relacional de las cooperativas de crédito en España: un estudio cualitativo de sus intangibles sociales mediante el análisis Delphi. *REVESCO Revista de Estudios Cooperativos* 101, 107-131.
<http://revistas.ucm.es/index.php/REVE/article/view/REVE1010230107A/18638>
- Seidman, S. A. y Zager, J. (1986-1987). The Teacher Burnout Scale. *Educational Research Quarterly*, 11(1), 26-33.
- Selye, H. (1936). A syndrome produced by diverse nocuous agents. *Nature*, 138, 32. doi: 10.1038/138032a0
- Selye, H. (1960). *La tensión en la vida*. Buenos Aires. Compañía General Fabril Editora (Orig. 1956).
- Selye, H. (1974). Stress without distress. En: Jesús Rodríguez Marín, editor. *Psicología Social de la Salud*. España; Editorial Síntesis Psicológica.
- Semmer, N. K., Grebner, S. y Elfering, A. (2004). Beyond self-report: using and situation-based Measures physiological, observational, in research on occupational stress. En P. Perrewé y D. Ganster (Eds.), *Emotional and physiological processes and positive intervention strategies. Research in occupational stress and well-being* (pp. 205-263). Boston: Elsevier Ltd.
- Sennett, R. (2006). *La cultura del nuevo capitalismo*. Barcelona: Anagrama.
- Seo, D. C. (2005). An explicative model of unsafe work behavior. *Safety Science*, 43, 187-211.

- Serrano, M. A., Moya-Albiol, L. y Salvador, A. (2009). Estrés laboral y salud: indicadores cardiovasculares y endocrinos. *Anales de Psicología*, 25, 150-159.
- Sheiner, E. K., Sheiner, E., Carel, R., Potashnik, G. y Shoham-Vardi, I. (2002). Potential association between male infertility and occupational psychological stress. *Journal of Occupational and Environmental Medicine*, 44(12), 1093-1099.
- Shima, S. y Satoh, E. (2006). Somatoform disorders in the workplace in Japan. *Int Rev Psychiatry*, 18(1), 35-40.
- SHL (1999). *OPQ32 Manual and User's Guide*. Surrey: SHL.
- Siegrist, J. (1996a). Adverse health effects of high effort /low reward conditions. *Journal of Occupational Health Psychology*, 1, 27-41.
- Siegrist, J. (1996b). *Stressful work, self-experience, and cardiovascular disease prevention*. En K. Orth-Gomér y N. Schneiderman (Eds.), *Behavioral Medicine Approaches to Cardiovascular Disease Prevention*. Mahwah, New Jersey: LEA.
- Siegrist, J. (1998). Adverse health effects of effort-reward imbalance at work: theory, empirical support and implications for prevention. En C. L. Cooper (Ed.), *Theories of Organizational Stress* (pp. 190-204). Oxford: Oxford University Press.
- Siegrist, J. (2002). Effort-Reward Imbalance at Work and Health. En P. L. Perrewé y D. C. Ganster. (Eds), *Historical and Current Perspectives on Stress and Health* (261-291). Oxford: Elsevier Science.
- Siegrist, J., Peter, R., Cremer, P. y Seidel, D. (1997). Chronic work stress is associated with atherogenic lipids and elevated fibrinogen in middle-aged men". *Journal of Internal Medicine*, 242(2), 149-156.
- Sigelman, C. K., Budd, E. C., Spanhel, C. L. y Schoenrock, C. J. (1981). Asking questions of retarded persons: A comparison of yes-no and either-or formats. *Applied Research in Mental Retardation*, 2, 347-357. doi: 10.1016/0270-3092(81)90029-1
- Silva, F. (1989). *Evaluación conductual y criterios psicométricos*. Madrid: Pirámide.

- Sim, J. y Wright, C. (2005) The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, 85(3), 257-268
- Sinclair, R., Sears, L. E., Probst, T. M. y Zajack, M. (2010). A multilevel model of economic stress and employee well-being. En J. Houdmont & S. Leka (Eds.) *Contemporary Occupational Health Psychology: Global Perspectives on Research and Practice* (vol.1, pp. 1-21). Hoboken, NJ: Wiley-Blackwell.
- Smith, S. T. (2005). On construct validity: Issues of method measurement. *Psychological Assessment*, 17, 396-408.
- Smyth, J. M., Soefer, M. H., Hurewitz, A., Kliment, A. y Stone, A. A. (1999). Daily psychosocial factors predict levels and diurnal cycles of asthma and peak flow. *Journal Behavioral Medicine* 22(2), 179-193.
- Solanes, A. y Zaragoza, B. M. (2010). ¿Trabajo de preferencia o contrato indefinido? Un análisis de su impacto sobre el estado del contrato psicológico. *Acciones e Investigaciones Sociales*, 28, 57-77.
- Sonnentag, S., Perrewé, P. L. y Ganster, D.C. (Eds.). (2009). *Research in Occupational stress and wellbeing. Volume 7: Current perspectives on job-stress recovery*. UK: Emerald/JAI.
- Soria, J. y Martínez, Rafael, Á. (2010). Desarrollo y validación de un instrumento de medida de las prácticas de gestión de la calidad. *Avances en Supervisión Educativa*, 12, 1-22. Recuperado de <https://avances.adide.org/index.php/ase/article/view/443>
- Spector, P. E. (1992). *Summated Rating Scale Construction*. Newbury Park, California: Sage Publications.
- Spector, P. E. (1994). Using self-report questionnaires in OB research: A comment on the use of a controversial method. *Journal of Organizational Behavior*, 15, 385-392.
- Stager, P., Hameluck, D. y Jubis, R. (1989). Underlying factors in air traffic control incidents. *Proceedings of the Human Factors Society 33rd Annual Meeting*, 43-46. Santa Monica, CA: Human Factors Society.

- Stansfeld, S. A., Bosman, H., Hemingway, H. y Marmot, M. G. (1998). Psychosocial work characteristics and social support as predictors of SF-36 health functioning: The Whitehall II Study. *Psychosomatic Medicine*, 60, 247-255.
- Stansfeld, S. A. y Candy, B. (2006). Psychosocial work environment and mental health – a meta-analytic review. *Scandinavian Journal of Work, Environment and Health* 32(6), 443–462.
- Stansfeld, S. A., Clark, C., Caldwell, T., Rodgers, B. y Power, C. (2008). Psychosocial work characteristics and anxiety and depressive disorders in midlife: the effects of prior psychological distress. *Occupational and Environmental Medicine*, 65(9), 634–642. doi: 10.1136/oem.2007.036640
- Stansfeld, S. A., Fuhrer, R., Shipley, M. J. y Marmot, M. G. (1999). Work characteristics predict psychiatric disorder: prospective results from the Whitehall II study *Occupational and Environmental Medicine*, 56, 302-307.
- Stansfield, S y Marmot, M. (Eds.). (2002). *Stress and the heart. Psychosocial pathways to coronary heart disease*. London: BMJ Books.
- Staw, B. M., Bell, N. E. y Clausen, J. A. (1986). The dispositional approach to job attitudes: A lifetime longitudinal test. *Administrative Science Quarterly*, 31, 56-77.
- Staw, B.M., Sandelands, L. E. y Dutton J. E. (1981). Threat-rigidity effects in organizational behavior: A multilevel analysis. *Administrative Science Quarterly*, 26(4), 501.
- Steenland, K., Fine, L., Belkic, K., Landsbergis, P., Schnall, P., Baker, D,... Tuchsén, F. (2000). Research findings linking workplace factors to CVD outcomes. *Occupational Medicine*, 15(1) 7-68.
- Steiger, J. H. y Lind, J. (1980). Statistically based tests for the number of common factors. Comunicación presentada en el *meeting anual de la Psychometric Society*. Iowa City, Mayo de 1980
- Stening, B. W. y Everett, J. E. (1984). Response styles in a cross-cultural managerial study. *Journal of Social Psychology*, 122, 151-156.

- Stephens, A., Lipsey, Z. y Wardle, J. (1998). Stress, hassles and variations in alcohol consumption, food choice and physical exercise: A diary study. *British Journal of Health Psychology*, 3, 51–63.
- Stöber, J., Dette, D. y Musch, J. (2002). Comparing Continuous and Dichotomous Scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment*, 78, 370–389.
- Strauss, H. J. y Zeigler, L. H. (1975). Delphi technique and its uses in social science research. *Journal of Creative Behavior*, 9(4), 253–259.
- Stricker, L. J. (2013). *ETS Research on Cognitive, Personality, and Social Psychology: I* (Research Report 13–01). Princeton, NJ: Educational Testing Service.
- Suls, J. (2001). Affect, stress and personality. En J. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 392–209). Mahwah, NJ: Erlbaum.
- Sutherland, V. y Cooper, C. (1988). Sources of Work Stress. En Hurrell, Murphy, Sauter and Cooper (eds.), *Occupational Stress: Issues and Developments in Research*. London, Taylor and Francis.
- Swearingen, D. L. (1997). Response sets, item format, and thinking style: Implications for questionnaire design. *Dissertation Abstracts International*, 98, 04094.
- Swearingen, D. L. (1998). Extreme responding style and the concreteness-abstractness dimension. *Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA*. (ERIC Document Reproduction No. ED423298). Recuperado de <http://files.eric.ed.gov/fulltext/ED423298.pdf>
- Takala, J. (2010, junio). Expectativas sobre la Estrategia Comunitaria de Salud y Seguridad en el Trabajo. *Conferencia europea de contribución a la evaluación a medio término de la estrategia comunitaria de salud y seguridad en el trabajo 2007-2012*. Barcelona.
- Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. En K. A. Bollen y J. S. Long (Eds.), *Testing structural equation models* (pp. 10–40). Newbury Park: Sage.

- Tanaka, J. S. y Huba, G. J. (1985). A fit index for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, 38, 197-201.
- Ten Berge, J. M. F. y Kiers, H. A. L. (1991). A numerical approach to the approximate and the exact minimum rank of a covariance matrix. *Psychometrika*, 56 (2), 309-315.
- Ten Brummelhuis, L. L., Bakker, A. B., Hetland, J. y Keulemans, L. (2012). Do new ways of working foster work engagement? *Psicothema*, 24(1), 113-120. Recuperado de <http://www.redalyc.org/articulo.oa?id=72723431018>
- Tena, G. (2002). El contrato psicológico: relación laboral empresa-trabajador. *Acciones e investigaciones sociales*, 15, 85-107.
- Theorell T. (1999). How to deal with stress in organizations?-A health perspective on theory and practice. *Scandinavian journal of work, environment & health*, 25(6), 616-624.
- Theorell T. (2000). Psychosocial factors at work in relation to musculoskeletal conditions. Implications for job design and rehabilitation. En Violante, F., Armstrong, T. y Kilbom A. (Ed.) *Occupational ergonomics. Work related musculoskeletal disorders of the upper limb and back*. New York: Taylor & Francis.
- Theorell, T. (2012). Evaluating life events and chronic stressors in relation to health: Stressors and health in clinical work. *Advances in psychosomatic medicine*, 32, 58-71.
- Theorell, T. y Karasek, R. A. (1996). Current issues relating psychosocial job strain and cardiovascular disease research. *Journal of Occupational Health Psychology*, 1(1), 9-26.
- Thissen, D., Steinberg, L. y Gerrard, M. (1986). Beyond Group-Mean Differences. The Concept of Item Bias. *Psychological Bulletin*, 99(1), 118-128. doi: 10.1037/0033-2909.99.1.118
- Thomas, D. R. y Zumbo, B. D. (1996). Variable importance in regression and related analysis. *Paper presented at the Annual Meeting of the Psychometric Society*, Banff, AB, Canada.
- Thorndike, E. L. (1920). A constant error on psychological rating. *Journal of Applied Psychology*, IV, 25-29.
- Thurstone, L. L. (1928). Attitudes Can Be Measured. *American Journal of Sociology* 33, 529-554.

- Tobías, A. (1999). Validez y sesgos en la investigación médica. Notas de metodología estadística. *JANO*, 56(1297), 76-77.
- Tolon, K. (2012). Futures Studies: A New Social Science Rooted in Cold War Strategic Thinking. En Solovey and Cravens (Eds.), *Cold War Social Science: Knowledge Production, Liberal Democracy, and Human Nature*, (pp. 45–62). New York: Palgrave Macmillan.
- Tomás, J. M., Galiana, L., Hontangas, P., Oliver, A. y Sancho, P. (2013). Evidencia acumulada sobre los efectos de método asociados a ítems invertidos. *Psicológica* 34, 365-381.
- Tomás, J. M., Rodrigo, M. y Oliver, A. (2005). Modelos lineales y no lineales en la explicación de la siniestralidad laboral. *Psicothema*, 17(1), 154-163.
- Tomás, J. M., Sancho, P., Oliver, A., Galiana, L. y Meléndez, J. C. (2012). Efectos de método asociados a ítems invertidos vs. ítems en negativo. *Revista Mexicana de Psicología*, 29(2), 105-115.
- Tomasina, F. (2012). Los problemas en el mundo del trabajo y su impacto en salud. Crisis financiera actual. *Revista de salud pública*, 14(1), 56-67.
- Torp, S., Riise, T. y Moen, B. E. (2001). The impact of psychosocial work factors on musculoskeletal pain: A prospective study. *Journal of Occupational & Environmental Medicine*, 43(2), 120–126.
- Torres, S. J. y Nowson, C. A. (2007). Relationship between stress, eating behavior, and obesity. *Nutrition*, 23, 887–894.
- Tourangeau, R. y Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, 103, 299-314.
- Tous-Pallarès, J. Bonasa-Jiménez, M. P. Mayor-Sánchez, C. y Espinoza-Díaz, I. M. (2011). Escala de Clima Psicosocial en el Trabajo: desarrollo y validación. *Anuario de Psicología*, 34(1), 214-226.
- Towers, A. M., Kishchuck, N., Sylvestre, M., Peters, C. y Bourgault, C. A. (1994). A qualitative investigation of organizational issues in an alcohol awareness program for blue-collar workers. *American Journal of Health Promotion*, 9(1), 56-63. doi: 10.4278/0890-1171-9.1.56

- Trice, H. M. y Roman, P. M. (1978). *Spirits and demons at work: Alcohol and other drugs on the job*. Ithaca, NY: Cornell University, New York State School of Industrial and Labor Relations.
- Trillo, M. (2009). Los contratos temporales y a tiempo parciales como formas de trabajo atípicas. Especial referencia a su régimen de seguridad social. *Revista Gaceta Laboral*, 15(1), 101-130.
- Trujillo, M. M., Valderrabano, M. de la L. y Hernández, R. (2007). Mobbing: historia, causas, efectos y propuesta de un modelo para las organizaciones mexicanas. *Revista Innovar*, 17(29), 71-91.
- Tsutsumi, A., Kayaba, K., Theorell, T. y Siegrist, J. (2001). Association between job stress and depression among Japanese employees threatened by job loss in a comparison between two complementary job-stress models. *Scandinavian Journal of Work, Environment & Health*, 27(2), 146-153.
- Tsutsumi, A., Kayaba, K., Ojima, T., Ishikawa, S. y Kawakami, N. (2007). Low control at work and the risk of suicide in Japanese men: a prospective cohort study. *Psychotherapy and Psychosomatics*, 76(3), 177-185.
- Tucker, L. R. (1951). *A method for synthesis of factor analysis studies*. Personnel Research Section Report, 984. Washington, D. C.: Department of the Army.
- Turoff, M. (1970). The design of a policy Delphi. *Technological Forecasting and Social Change*, 2(2), 149-171.
- UGT (2013). *Costes Socio-Económicos de los Riesgos Psicosociales*. Observatorio de Riesgos Psicosociales UGT-CEC.
- Undén, A. L. (1996). Social Support at Work and Its Relationship to Absenteeism. *Work Stress*, 10(1), 46-61.
- University of Melbourne (2013). *Workplace prevention of mental health problems: Guidelines for organisations*. Melbourne School of Population and Global Health: University of Melbourne

- Useche, M. C. (2002). La organización del trabajo en el marco de la globalización. *Gaceta Laboral*, 8, 67-78.
- Usha, P. (2016). A comparative study to assess levels of stress between general ward and OT nurses at selected hospitals, chittoor district. *International Journal of Recent Scientific Research*, 7(4), 10140- 10143. Recupeado de <http://www.recentscientific.com/sites/default/files/4927.pdf>
- Vagg, P. R. y Spielberger, C. D. (1999). The Job Stress Survey: Assessing perceived severity and frequency of occurrence of generic sources of stress in the workplace. *Journal of Occupational Health Psychology*, 4(3), 288–292.
- Van Der Doef, M. y Maes, S. (1999). The Job Demand- Control (Support) Model and psychological well-being: A review of 20 years of empirical research. *Work and Stress*, 13(2), 87-114.
- Van Dick, R. y Wagner, U. (2001). Stress in teaching. A structural equation approach. *British Journal of Educational psychology*, 71, 243-259.
- Van Herk, H., Poortinga, Y. H. y Verhallen, T. M. M. (2004). Response styles in rating scales: Evidence of method bias in data from 6 EU countries. *Journal of Cross-Cultural Psychology*, 35(3), 346-360.
- Van Laar, D., Edwards, J. y Easton, S. (2007). The Work-Related Quality of Life scale for healthcare workers. *Journal of Advanced Nursing*, 60(3), 325–333.
- Van Veldhoven, M. van, Meijman, T. F., Broersen, J. P. J. y Fortuin, R. J. (1997). *VBBA Test manual*. Amsterdam: Stichting Kwaliteitsbevordering Bedrijfsgezondheidszorg.
- Varela-Ruiz, M., Diaz-Bravo, L. y García-Durán, R. (2012). Descripción y usos del método Delphi en investigaciones del área de la salud. *Revista de Investigación en Educación Médica*, 1(2), 90-95.
- Vega, S. (2003). *Riesgo psicosocial el modelo demanda- control apoyo social (I)*. (NTP-603). Barcelona: INSHT.
- Vernon, P. E. (1934). The attitude of the subject in personality testing. *Journal of Applied Psychology*, 18, 165-177.

- Verona, M. C. y Santana, R. (2012). El mobbing: ¿el crimen perfecto? No, qué más quisieran “algunos.” *Criterio Libre*, 10(17), 147–176.
- Vesga, J. J. (2007). Contenido del contrato psicológico percibido en distintas modalidades de trabajo. *Perspectivas en Psicología*, 10, 79-99.
- Vesga J. J. (2011). Los tipos de contratación laboral y sus implicaciones en el contrato psicológico. *Pensamiento Psicológico*, 9(16), 171-181.
- Vigil-Colet, A., Ruiz-Pamies, M., Anguiano-Carrasco, C. y Lorenzo-Seva, U. (2012). The impact of social desirability on psychometric measures of aggression. *Psicothema*, 24(2), 310–315. doi:10.7334/psicothema2012.297
- Viladrich, C. y Doval, E. (2011). *Medición, fiabilidad y validez* (6ªed.). Bellaterra: Laboratori d'Estadística aplicada i Modelització (UAB).
- Villalobos, G. (2004). Vigilancia Epidemiológica de los Factores Psicosociales. Aproximación Conceptual y Valorativa. *Ciencia & Trabajo*, 6(14), 197–201.
- Virtanen, M., Kivimäki, M., Joensuu, M., Virtanen, P., Elovainio, M. y Vahtera J. (2005). Temporary employment and health: a review. *International Journal of Epidemiology*, 34(3), 610-622.
- Viswanathan, M. (2005). *Measurement error and research design*. Thousand Oaks, CA: Sage. doi: org/10.4135/9781412984935
- Viswesvaran, C., Sánchez, J. I. y Fisher, J. (1999). The role of social support in the process of work stress: A meta-analysis. *Journal of Vocational Behavior*, 54(2), 314-334.
- Vivas, V., Delgadillo., L. G. y Mercado, A. (2008). Violencia laboral, su significación desde hombres y mujeres. *Ciencia & Trabajo*, 30, 151-156.
- Von der Gracht, H. A. (2012). Consensus measurement in Delphi studies: Review and implications for future quality assurance. *Technological Forecasting and social change* 79(8), 1525–1536.
- Vrijkotte, T. G., van Doornen, L. J. y de Geus, E. J (2000). Effects of Works stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension* 35(4), 880-886.

- Wallis, D. J. y Hetherington, M. M. (2009). Emotions and eating. Self-reported and experimentally induced changes in food intake under stress. *Appetite*, 52(2), 355-362.
- Wang, J. L. (2008). The relationship between work stress and mental disorders in men and women: findings from a population-based study. *Journal Epidemiology Community Health*, 62(1), 42-47.
- Wang, J. L. y Patten, S. B. (2001). Perceived work stress and major depression in the Canadian employed population, 20-49 years old. *Journal of Occupational Health Psychology*, 6, 283-289.
- Wang, J. L., Patten, S. B., Currie, S., Sareen, J. y Schmitz, N. (2012). A population-based longitudinal study on work environmental factors and the risk of major depressive disorder. *American Journal Epidemiology*, 176, 52-59.
- Ward, A. y Mann, T. (2000). Don't Mind If I Do: Disinhibited Eating Under Cognitive Load. *Journal of Personality and Social Psychology*, 78, 753-763.
- Ware, J. E. Jr. (1978), Effects of Acquiescent Response Set on Patient Satisfaction Ratings. *Medical Care*, 16, 327-336.
- Warr, P. (1987). *Work, Unemployment and Mental Health*. Oxford: Clarendon Press.
- Warr, P. (1994). A conceptual frame work for the study of work and mental health. *Work & Stress*, 8, 84-97.
- Warr, P. (1996). Employee well-being. En P.B. Warr (Ed.), *Psychology at work* (4th edition, pp. 224-253). Harmondsworth: Penguin.
- Warr, P. (1998). Well-being and the workplace. En D. Kahneman, E. Diener y N. Schwarz (Eds), *Understanding Quality of Life: Perspectives on Enjoyment and Suffering*. Nueva York: Russell Sage.
- Warr, P. (2007). *Work, Happiness, and Unhappiness*. Lawrence Erlbaum Associates, Inc., Publishers: New Jersey.
- Warr, P., Cook, J. y Wall, T. (1979). Scales for the measurement of some work attitudes and aspects of psychological well-being. *Journal of Occupational Psychology*, 52, 129-148.

- Watson, D., Clark, L.A. y Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Watson, D. y Pennebaker, J. W. (1989). Health complaints, stress and distress: Exploring the central role of negative affectivity. *Psychological Review*, 96(2), 234-254.
- Watson, D. y Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98, 219-235.
- Weinberg, A., Cooper, C. L. y Weinberg, A. (1999). Workload, stress and family life in British Members of Parliament and the psychological impact of reforms to their working hours. *Stress Medicine*, 15(2), 79-87.
- Weijters, B., Geuens, M. y Schillewaert, N. (2010). The individual consistency of acquiescence and extreme response style in self-report questionnaires. *Applied Psychological Measurement*, 34(2), 105-121.
- Weiss, H. M. y Cropanzano, R. (1996). Affective Events Theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Research in Organizational Behavior*, 18, 1-74.
- West, S. G., Finch, J. F. y Curran, P. J. (1995). Structural Equation Models with non-normal variables - problems and remedies. En R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA: Sage.
- Wheatley D. (1990). The stress profile. *British Journal of Psychiatry*, 156, 685-688.
- WHO Regional Office for Europe (2010). *Mental Health and Wellbeing at the Workplace – Protection and Inclusion in Challenging Times*. WHO Regional Office for Europe, Copenhagen.
- Wieclaw, J., Agerbo, E., Mortensen, P. B., Burr, H., Tuchsén, F. y Bonde, J. P. (2008). Psychosocial working conditions and the risk of depression and anxiety disorders in the Danish workforce. *BMC Public Health*, 8, 280-288. doi: 10.1186/1471-2458
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, 1, 80-83.

- Wilhelm, K., Kovess, V., Rios-Seidel, C. y Finch, A. (2004). Work and mental health. *Social Psychiatry & Psychiatric Epidemiology*, 39(11), 866–873.
- Williams, C. (ed.). (1998). *Social Factors, Work, Stress and Cardiovascular Disease*. Prevention in the European Union. Brussels: The European Heart Network.
- Williams, R. B. (2003). Invited commentary: socioeconomic status, hostility, and health behaviors--does it matter which comes first? *The American Journal of Epidemiology*, 158(8), 743-760.
- Williams, A. y Cooper, C. (1998). Measuring occupational stress: Development of the Pressure Management Indicator. *Journal of Occupational Health Psychology* 3, 306-321.
- Winocur, E., Uziel, N., Lisha, T., Goldsmith, C. y Eli, I. (2011). Self-reported bruxism - associations with perceived stress, motivation for control, dental anxiety and gagging. *Journal of Oral Rehabilitation*, 38(1), 3-11. doi:10.1111/j.1365-2842.2010.02118.x.
- Woo, J. M. y Postolache, T. T. (2008). The impact of work environment of mood disorders and suicide: evidence and implications. *International journal on disability and human development*, 7, 185–200.
- Woudenbergh, F. (1991). An evaluation of Delphi. *Technological Forecasting and Social Change*, 40(2), 131-150.
- Wright, J. D. (1975). Does Acquiescence Bias the “Index of Political Efficacy? *Public Opinion Quarterly*, 39(2), 219-226.
- Yanes, J. (2008). *El control de estrés y el mecanismo del miedo*. Madrid. Editorial Edaf.
- Yankelevich, M., Broadfoot, A., Gillespie, J. Z., Gillespie, M. A. y Guidroz, A. (2012). General Job Stress: A unidimensional measure and its non-linear relations with outcome variables. *Stress and Health*, 28(2), 137-148.
- Yañez, R. y Cuadra, R. (2008) La técnica Delphi y la investigación en los servicios de salud. *Ciencia y Enfermería*, XIV(1), 9-15. doi:10.4067/S0717-95532008000100002
- Yela, M. (1968). *Apuntes de psicometría y estadística*. Madrid: Universidad Complutense de Madrid.

- Yela, M., (1987). *Introducción a la teoría de los tests* (pp. 3-29). Madrid.
- Ylipaavalniemi, J., Kivimäki, M., Elovainio, M., Virtanen, M., Keltikangas-Järvinen, L. y Vahtera, J. (2005). Psychosocial work characteristics and incidence of newly diagnosed depression: A prospective cohort study of three different models. *Social Science and Medicine*, 61(1), 111-122.
- Ylitalo, J. (2009). *Controlling for common method variance with partial least squares path modeling: A Monte Carlo study* (Mat-2.4108 Independent Research Projects in Applied Mathematics 64353A). Helsinki: Helsinki University of Technology, Department of Mathematics and Systems Analysis. Recuperado de http://salserver.org.aalto.fi/vanhat_si_vut/Opinnot/Mat-2.4108/pdf-files/eyli09b.pdf
- Zárraga, C. L., Molina, M. V. M. y Corona, S. E. (2013). Estudio de caso: Análisis de la aplicación de la segmentación de mercado como estrategia para las pequeñas empresas. *Revista Internacional Administración & Finanzas*, 6(5) 109-119.
- Zartha, J. W., Montes, J.M., Toro, I. D. y Villada, H. S. (2014) Método Delphi - Propuesta para el cálculo del número de expertos en un estudio Delphi sobre empaques biodegradables al 2032. *Revista Espacios*, 35(13), 10-11.
- Zavala, A. (1965). Development of the forced-choice rating scale technique. *Psychological Bulletin*, 63(2), 117-124
- Zautra, A. J, Fasman, R., Reich, J. W., Harakas, P., Johnson, L. M., Olmsted, M. E. y Davis, M. C. (2005). Fibromyalgia: evidence for deficits in positive affective regulation. *Psychosomatic Medicine*, 67, 147-155.
- Zhang, X. y Savalei, V. (2015). Improving the factor structure of psychological scales: The Expanded format as an alternative to the Likert scale format. *Educational and Psychological Measurement*, 76(3), 357-386. doi:10.1177/0013164415596421
- Zhang, X., Noor, R. y Savalei, V. (2016). Examining the Effect of Reverse Worded Items on the Factor Structure of the Need for Cognition Scale. *PLOS ONE* 11(6), 1-15. doi:10.1371/journal.pone.0157795

- Zohar, D. (1997). Predicting burnout with a hassle-based measure of role demands. *Journal of Organizational Behavior*, 18(2), 101-115.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of Differential Item Functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) items scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. En C. R. Rao y S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics* (pp. 45-79). Amsterdam, Netherlands: Elsevier Science B.V. doi: 10.1016/S0169-7161(06)26003-6





INDICE TABLAS

Tabla 1: Factores de Riesgo Psicosocial en el Trabajo	13
Tabla 2: Riesgos Psicosociales.....	14
Tabla 3: Escala de estrés de Holmes y Rahe	18
Tabla 4: Trastornos asociados con el estrés	20
Tabla 5: Trastornos asociados al Burnout	25
Tabla 6: Trastornos asociados al Mobbing.	28
Tabla 7: Comparación de porcentajes de exposición a factores psicosociales entre ENCT VI y ENCT VII	30
Tabla 8: Ranking de los factores de riesgos incluidos en las evaluaciones de riesgos laborales	32
Tabla 9: Riesgo psicosocial enmarcado dentro de la Ley de Prevención de Riesgos Laborales.....	35
Tabla 10: Técnicas de evaluación de los factores psicosociales dependiendo del objetivo del análisis	38
Tabla 11: Clasificación test psicométricos.....	45
Tabla 12: Instrumentos de investigación de riesgos psicosociales según información aportada	51
Tabla 13: Revisión FPSICO 3.1 (2014)	54
Tabla 14: Revisión Cuestionario Multidimensional DECORE.....	55
Tabla 15: Revisión CoPsoQ-ISTAS 21.....	56
Tabla 16: Revisión INERMAP	57
Tabla 17: Revisión MC Mutual-UB.....	58
Tabla 18: Revisión Prevenlab Psicosocial	59
Tabla 19: Revisión MARC-UV.....	60
Tabla 20: Revisión WONT Fuente: Elaboración propia a partir de Moreno y Báez (2010) y CEOE (2013).	62
Tabla 21: Criterios de selección de la herramienta adecuada.....	63
Tabla 22: Sesgos y Fuentes de error en escalas de autoinforme	69
Tabla 23: Catálogo de sesgos clasificados por fuentes	71
Tabla 24: Escalas de DS.....	81
Tabla 25: Fuentes de Varianza Común de Método	85
Tabla 26: Correspondencia entre ítems del cuestionario y proposiciones derivadas marco teórico	98
Tabla 27: Valoraciones para el cálculo del Coeficiente de conocimiento de experto	102
Tabla 28: Calendario de envío y recepción de cuestionarios	104
Tabla 29: Composición del panel de 16 expertos en materia de riesgos psicosociales.....	105
Tabla 30: Estadísticos 1ª Ronda (1R).....	106
Tabla 31: Criterios para considerar la falta de consenso en los ítems.....	108
Tabla 32: Grado de acuerdo alcanzado según la concordancia de opiniones.....	108
Tabla 33: Estadísticos de medida de estabilidad grupal.....	111

Tabla 34: Valores de Kappa	112
Tabla 35: Porcentaje de acuerdo entre panelistas R1 (N=16).....	113
Tabla 36: Resultados estadísticos de la primera ola: Criterios de no consenso.....	113
Tabla 37: Resultados estadísticos de la primera ola: Grado de consenso.....	114
Tabla 38: Porcentaje de acuerdo entre panelistas R2 (N=16).....	115
Tabla 39: Resultados estadísticos de la segunda ola: Criterios de no consenso	115
Tabla 40: Ítems sin consenso 2ª vuelta	116
Tabla 41: Resultados estadísticos de la segunda ola: Grado de consenso	116
Tabla 42: Criterios para considerar los acuerdos como significativos	118
Tabla 43: Resultados ítems cualitativos: Valoraciones ítem A10a.....	119
Tabla 44: Resultados ítems cualitativos: Valoraciones ítem A17a.....	119
Tabla 45: Valoraciones del apartado I: “Afirmaciones generales”	120
Tabla 46: Valoraciones del apartado II: “Diseño de instrumentos de medida”	121
Tabla 47: Valoraciones del apartado III: “Aplicación”	122
Tabla 48: Valoraciones del apartado IV: “Corrección e interpretación de resultados”	123
Tabla 49: Resultados ítems cualitativos: Valoraciones comentarios finales	123
Tabla 50: Resultados de los medidores de estabilidad	125
Tabla 51: Principales procedimientos y directrices sobre la evaluación de riesgos psicosociales	136
Tabla 52: Fases de la depuración de datos	148
Tabla 53: Muestreo en base al número de ocupados por grupo de sexo	149
Tabla 54: Distribución en función del sexo y de la edad.....	150
Tabla 55: Distribución en función de la nacionalidad.....	150
Tabla 56: Distribución según la situación laboral	151
Tabla 57: Distribución según el puesto de trabajo.....	151
Tabla 58: Distribución según la antigüedad en el puesto y en la empresa	152
Tabla 59: Distribución en función de la baja médica	152
Tabla 60: Marcadores de deseabilidad social.	157
Tabla 61: Fiabilidad de las escalas de medida.....	158
Tabla 62: Criterios de magnitud de DIF	162
Tabla 63: Estadísticos descriptivos de los ítems del CoPsoQ-ISTAS21 y de DS (n1=325)	164
Tabla 64: Índices de bondad de ajuste del modelo sugerido por el AFE.....	169
Tabla 65: Matriz de cargas factoriales del AFE controlando DS (n=563)	170
Tabla 66: Puntuaciones factoriales para los grupos bajos y altos en DS (n=563)	171
Tabla 67: Funcionamiento diferencial del ítem en AN (n=563).....	172
Tabla 68: Funcionamiento diferencial del ítem purificado en AN (n=563)	173
Tabla 69: Puntuaciones factoriales para los grupos bajos y altos en AN (n=563)	174

INDICE FIGURAS

Figura 1: Relación factores de riesgo psicosocial – enfermedad	16
Figura 2: Fases evaluación factores psicosociales	37
Figura 3: Modelo Demanda- Control.....	41
Figura 4: Modelo Esfuerzo-Recompensa.....	42
Figura 5: Errores de medida.....	65
Figura 6: Diferentes errores y sesgos en las diferentes etapas del proceso investigador	66
Figura 7: Fases del método Delphi.....	97
Figura 8: Comparación entre rondas del resultado global.....	117
Figura 9: Comparación RIR entre Rondas	125
Figura 10 Comparación χ^2 entre Rondas.....	126
Figura 11. Resultados del análisis paralelo (n=325).	166





ANEXO I
CUESTIONARIO PRIMER ESTUDIO





NOMBRE:

Tal y como se desprende del texto de presentación, he pensado en usted para formar parte del panel de expertos, y así poder someter a juicio determinadas conclusiones referentes a la evaluación de riesgos psicosociales en el trabajo.

Si acepta colaborar en el presente panel de expertos, se le pedirá que participe solamente en dos rondas. Debido a que el éxito de este estudio está basado en el cumplimiento de su compromiso de colaboración, es necesario que se respondan las preguntas del test dentro de los plazos indicados:

Plazo para recepción de cuestionarios: 20/01/2016.

Antes de comenzar con el primer cuestionario, agradeceríamos que realizase una autovaloración acerca de su conocimiento sobre la materia:

- I. Marque la casilla que, a su juicio, corresponde con su grado de conocimiento o información acerca de los métodos utilizados en evaluaciones de riesgos psicosociales

0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- II. En la tabla que a continuación se presenta, seleccione el grado de influencia que considera que han tenido las siguientes fuentes, sobre la formulación de sus criterios acerca de los métodos de evaluación utilizados en evaluaciones de riesgos psicosociales (Deberá responder a todas las fuentes)

	FUENTES	Grado de influencia de cada una de las fuentes en sus criterios		
		A (alto)	M (medio)	B (bajo)
1º	Análisis teóricos realizados por usted	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2º	Su experiencia en el tema	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3º	Trabajos de autores nacionales consultados	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4º	Trabajos de autores extranjeros consultados	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5º	Su propio conocimiento del estado del problema en el extranjero	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6º	Su formación	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

CUESTIONARIO DE LA PRIMERA RONDA:

A continuación, le presentamos algunas afirmaciones relacionadas con la evaluación de riesgos psicosociales en las organizaciones. Su tarea consiste en valorar, siguiendo la escala propuesta, cada uno de los enunciados presentados:

Totalmente en desacuerdo	En desacuerdo	Indeciso	De acuerdo	Totalmente de acuerdo
1	2	3	4	5

Afirmaciones generales		Valoración
A1	La mayoría de empresas concede más importancia a la evaluación de los riesgos físicos que a la evaluación de los riesgos psicosociales	
A2	Son pocas las empresas que cuentan con procedimientos para gestionar riesgos psicosociales tales como estrés, <i>mobbing</i> , <i>burnout</i> , violencia, etc.	
A3	A menudo se confunde lo que es un método de evaluación de riesgos (método científico), con el hecho de administrar cuestionarios que evalúan riesgos	
A4	La evaluación de riesgos implica más procesos que valorar los riesgos mediante un cuestionario	
Observaciones:		
Diseño de instrumentos de medida		Valoración
A5	Para valorar los riesgos psicosociales es esencial la autopercepción del trabajador	
Para que los datos recogidos sobre la autopercepción del trabajador sobre su experiencia y exposición a los riesgos psicosociales sea lo más objetiva posible, la herramienta a utilizar para la recogida de datos debería contener los siguientes criterios:		
A6	a) Que cumpla propiedades psicométricas de fiabilidad y validez	
A7	b) Que controlen los sesgos	
A8	c) Baremación y estandarización	
A9	d) La fundamentación teórica	
A10a	e) Otros (indique cuales):	A10b
A11	Es poco apropiado incluir en los cuestionarios de evaluación de riesgos psicosociales escalas cuya finalidad no es detectar riesgos, como son escalas	
Observaciones:		

Aplicación		Valoración
A12	Las medidas de autoinforme son susceptibles de ser falseadas	
A13	A la hora de aplicar un test suele darse más importancia a las propiedades psicométricas de fiabilidad y validez que a la detección de sesgos	
Los siguientes sesgos de respuesta son algunos de los que pueden atribuirse a escalas de autoinforme utilizadas en prevención de riesgos psicosociales :		
A14	a) aquiescencia	
A15	b) la deseabilidad social	
A16	c) el autoengaño o la simulación	
A17a	d) Otros:	A17b
A18	Un protocolo de administración donde, por ejemplo, los encargados de administrar los test lean las instrucciones pausada y claramente y traten de	
Observaciones:		
Corrección e Interpretación de resultados		Valoración
A19	A la hora de administrar test de evaluación de riesgos psicosociales no es frecuente realizar control de posibles sesgos o fuentes de error	
A20	Los sesgos son un problema importante pues podrían contaminar los resultados de medida de los test, pudiendo incluso invalidar toda la evaluación	
A21	Los sesgos de respuesta son posiblemente los de mayor incidencia en los cuestionarios de autoinforme	
A22	Actualmente apenas existen herramientas para evaluar riesgos psicosociales que contengan control de sesgos	
A23	A pesar de existir técnicas destinadas a detectar y minimizar los sesgos de respuesta a menudo son eludidos	
A24	La percepción de incumplimiento de contrato psicológico, por parte del trabajador, podría influir en el falseamiento consciente de las respuestas de escalas utilizadas en la evaluación de riesgos psicosociales	
A25	Es necesario tener conocimientos de psicometría para interpretar los resultados de las escalas de autoinforme utilizadas para valorar los riesgos	
Observaciones:		
Sugerencias, Comentarios y/u Observaciones:		

Gracias por su colaboración



ANEXO II
CUESTIONARIO SEGUNDO ESTUDIO





Cuestionario de Evaluación de Riesgos Psicosociales en el trabajo.

Instrucciones

Este cuestionario está diseñado para identificar y medir todas aquellas condiciones de trabajo del ámbito psicosocial que pueden representar un riesgo para la salud y el bienestar de las personas trabajadoras. Forma parte de un trabajo de investigación en proceso, perteneciente a un programa de doctorado en Psicología Social, Básica y Metodología de la Universidad de Santiago de Compostela.

Es un cuestionario **ANÓNIMO**, de respuesta **VOLUNTARIA**. Toda la información **será analizada por personal técnico sujeto al mantenimiento del secreto profesional** y utilizada exclusivamente para los fines que hemos descrito. Ni en la respuesta, ni en la forma de distribución y de recogida, ni en el informe de resultados podrá ser identificada ninguna persona de forma individualizada. Los cuestionarios se destruirán una vez se analicen sus datos.

La respuesta es **INDIVIDUAL**. Por eso te pedimos que respondas sinceramente cada una de las preguntas, sin debatir con nadie, y sigas las instrucciones de cada una de ellas.

La mayoría de preguntas tienen **varias opciones de respuesta y te pedimos que señales con una "X" la respuesta que consideres que describe mejor tu situación** (por ejemplo, escogiendo una sola opción entre las posibles respuestas: "*siempre / muchas veces / algunas veces / sólo alguna vez / nunca*"). En otras preguntas no se trata de marcar una opción, sino de responder con un número. Utiliza el espacio de la última página para cualquier comentario respecto a esta encuesta. Te pedimos que respondas sinceramente a cada una de las preguntas sin previa consulta ni debate con nadie y que sigas las instrucciones de cada pregunta para contestar.

Muchas gracias por tu colaboración.

I. En primer lugar, nos interesan algunos datos sobre ti:

a) Eres

- 1 ☐ Hombre
- 2 ☐ Mujer

b) ¿Qué edad tienes?

- 1 ☐ Menos de 26 años
- 2 ☐ Entre 26 y 35 años
- 3 ☐ Entre 36 y 45 años
- 4 ☐ Entre 46 y 55 años
- 5 ☐ Más de 55 años

c) ¿De qué país eres?

- 1 ☐ (especificalo) _____

d) ¿Qué estudios has completado?

- 1 ☐ Primarios
- 2 ☐ Secundarios
- 3 ☐ Formación Profesional (FP)
- 4 ☐ Universitarios

e) ¿Cuál es tu situación laboral?

- 1 ☐ Desempleado (*Salta a la pregunta 9*)
- 2 ☐ Contrato Fijo
- 3 ☐ Contrato Temporal
- 4 ☐ Otros (*especificalo*) _____

f) ¿Qué puesto de trabajo ocupas? (sólo para personas empleadas)

- 1 ☐ (*especificalo*) _____

g) ¿Cuánto tiempo llevas en el mismo puesto de trabajo? (sólo para personas empleadas)

- 1 ☐ (*especificalo*) _____

h) ¿Cuánto tiempo llevas en la misma empresa? (sólo para personas empleadas)

- 1 ☐ (*especificalo*) _____

i) ¿Has estado de baja en alguna ocasión durante los últimos 12 meses?

- 1 ☐ No
- 2 ☐ Sí (*especifica cuanto tiempo*) _____
- 3 ☐ Marca esta casilla con una X si la causa de la baja ha estado relacionada con el trabajo

**II. A continuación, las preguntas tratan sobre las exigencias y contenidos de tu trabajo:
Por favor, elige UNA SOLA RESPUESTA para cada pregunta (Marca con una X)**

¿Con qué frecuencia...?	Siempre	Muchas veces	A veces	Sólo alguna vez	Nunca
1. ¿La distribución de tareas es irregular y provoca que se te acumule el trabajo?					
2. ¿Tienes tiempo suficiente para hacer tu trabajo?					
3. ¿Hay momentos en los que necesitarías estar en la empresa y en casa a la vez?					
4. ¿Sientes que tu trabajo te ocupa tanto tiempo que perjudica a tus tareas doméstico-familiares?					
5. ¿Alguna vez has dicho algo malo de otra persona?					
6. ¿En el trabajo tienes que ocuparte de los problemas personales de otras personas?					
7. ¿Tienes que trabajar muy rápido?					
8. ¿Alguna vez has cogido alguna cosa que no era tuya?					
9. ¿Tienes mucha influencia sobre las decisiones que afectan a tu trabajo?					
10. ¿Tienes influencia sobre CÓMO realizas tu trabajo?					

III. Siguiendo con los contenidos de tu trabajo, escoge 1 sola RESPUESTA para cada una de las siguientes preguntas.

¿En qué medida...?	En gran medida	En buena medida	En cierta medida	En alguna medida	En ningún caso
11. ¿Tu trabajo, en general, es desgastador emocionalmente?					
12. ¿El ritmo de trabajo es alto durante toda la jornada?					
13. ¿Alguna vez te has aprovechado de alguien?					
14. ¿Tu trabajo permite que aprendas cosas nuevas?					
15. ¿Tu trabajo permite que apliques tus habilidades y conocimientos?					
16. ¿Tus tareas tienen sentido?					
17. ¿Las tareas que haces te parecen importantes?					

IV. Estas preguntas hacen referencia al grado de definición de tus tareas, escoge 1 SOLA RESPUESTA para cada una de las siguientes preguntas

¿En qué medida...?	En gran medida	En buena medida	En cierta medida	En alguna medida	En ningún caso
18. ¿Tu trabajo tiene objetivos claros?					
19. ¿Sabes exactamente qué se espera de ti en el trabajo?					
20. ¿Se te exigen cosas contradictorias en el trabajo?					
21. ¿Tienes que hacer tareas que tú crees que deberían hacerse de otra manera?					
22. ¿Alguna vez has dejado para mañana lo que debes hacer hoy					
23. ¿En tu empresa se te informa con suficiente antelación de decisiones importantes, cambios y					
24. ¿Recibes toda la información que necesitas para realizar bien tu trabajo?					

V. Estas preguntas hacen referencia a las preocupaciones sobre posibles cambios en tus condiciones de trabajo. Escoge 1 SOLA RESPUESTA para cada una de las siguientes preguntas

En estos momentos, estás preocupado o preocupada por ... :	En gran medida	En buena medida	En cierta medida	En alguna medida	En ningún caso
25.... si te cambian el horario (turno, días de la semana, horas de entrada y salida) contra tu voluntad?					
26....si te varían el salario (que no te lo actualicen, que te lo bajen, que introduzcan el salario variable, que te paguen en especies, etc.)?					
27... si te despiden o no te renuevan el contrato?					
28....lo difícil que sería encontrar otro trabajo en el caso de que te quedaras en paro?					

VI. Las siguientes preguntas hacen referencia al reconocimiento, la confianza y la justicia en tu lugar de trabajo. Escoge 1 SOLA RESPUESTA para cada una de las siguientes preguntas

¿En qué medida...?	En gran medida	En buena medida	En cierta medida	En alguna medida	En ningún caso
29. ¿Confía la Dirección en que los trabajadores hagan un buen trabajo?					
30. ¿Te puedes fiar de la información procedente de la Dirección?					
31. ¿Se solucionan los conflictos de una manera justa?					
32. ¿Se distribuyen las tareas de forma justa?					
33. ¿Se puede afirmar que tu jefe inmediato planifica bien el trabajo?					
34. ¿Se puede afirmar que tu jefe inmediato resuelve bien los conflictos?					

VII. A continuación señale en qué medida está de acuerdo con los siguientes enunciados.

	Muy en desacuerdo	Algo en desacuerdo	En parte de acuerdo, en parte en desacuerdo	Algo de acuerdo	Muy de acuerdo
35. Me complace saber que mi trabajo ha contribuido al bien de la empresa					
36. Me siento parte de la empresa					
37. Incluso si esta organización no marcha bien, sería reacio/a a cambiar de organización					
38. En mi trabajo, me gusta sentir que estoy esforzándome no sólo por mi, sino también por la organización					
39. Estoy orgulloso/a de decirle a la gente la empresa en la que trabajo					
40. Una mejor oferta salarial no me haría considerar cambiar de trabajo					

VII.Las siguientes preguntas consisten en una serie de palabras que describen diferentes sentimientos y emociones.

Lee cada palabra y marca con una X el espacio correspondiente a la respuesta apropiada para ti. Indica cómo te sientes HABITUALMENTE EN EL TRABAJO (EN GENERAL). Por favor, responde a todas las preguntas y elige UNA SOLA RESPUESTA para cada una de ellas

	Nada	Poco	Algo	Bastante	Mucho		Nada	Poco	Algo	Bastante	Mucho
Interesado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Irritable	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Tenso/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Despierto/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Animado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Avergonzado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Disgustado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Inspirado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Enérgico/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Nervioso/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Culpable	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Decidido/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Asustado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Atento/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Enojado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Intranquilo/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Entusiasmado/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Activo/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Orgullosa/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	Temeroso/a	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>

Observaciones

¡MUCHAS GRACIAS POR TU TIEMPO Y COLABORACIÓN!